THE UNIVERSITY *of* EDINBURGH

LEXICAL CONDITIONS ON SYNTACTIC KNOWLEDGE: AUXILIARY SELECTION IN NATIVE AND NON-NATIVE GRAMMARS OF ITALIAN

ANTONELLA SORACE

PH.D
UNIVERSITY OF EDINBURGH
1992

A Luisa e Carmelo

DECLARATION

I hereby declare that this thesis is my original work and of my own execution and authorship.

ACKNOWLEDGMENTS

In developing the ideas presented in this thesis, I have had the benefit of interacting with many people who provided suggestions, clarification, and criticism. I am particularly indebted to my supervisor Ellen Gurman Bard for her critical and constructive support over the years: she inspired my interest in magnitude estimation of linguistic acceptability, constantly reminded me to bring theory back down to earth, and taught me that statistics can be fun. I am grateful to Wolfgang Klein for the useful discussions I had with him in the early stages of this research. Keith Mitchell and Bob Ladd shared their insights with me and gave me invaluable advice. Among the people with whom I discussed aspects of this work, I particularly wish to thank David Birdsong, Anne Cutler, Alan Davies, Elisabet Engdahl, Sue Gass, Eric Kellerman, Pim Levelt, Jim Miller, Lorenzo Renzi, Dan Robertson, Bill Rutherford, Jackie Schachter, Herbert Schriefers, Bonnie Schwartz, Mike Sharwood Smith, Edward Tuttle, Nigel Vincent, Lydia White, and Helmut Zobl.

I also would like to thank the audiences at the conferences where I presented parts of this work: three Language Acquisition Research Symposia (Utrecht 1986, 1988, 1992); two Second Language Research Forum conferences (Los Angeles 1987, 1991); two AILA World Congresses (Sydney 1987, Thessaloniki 1991); three meetings of the Linguistic Association of Great Britain (Durham 1988, Oxford 1991; York 1991); the meeting of the Association for French Language Studies, Edinburgh 1991; the Michigan State conference on Theory and Methodology in Second Language Acquisition, East Lansing 1991; the Boston University Conference on Language Development,1991.

I am grateful to the institutions, both in Scotland and in Italy, which gave me access to their students, and particularly to the many people who volunteered to participate in this study as informants.

Finally, this thesis could not have been written without the love, patience and support of my family: very special thanks to Marco, Andrew, and Bob.

ABSTRACT

This study has both a theoretical and a methodological dimension. Theoretically, it is concerned with variation and indeterminacy in linguistic acceptability judgments. Methodologically, it involves the application of a rigorous procedure for the elicitation of judgment data that is sensitive to informants' variable or indeterminate intuitions.

The theoretical focus is on the linguistic intuitions of native and non-native speakers of Italian about a number of grammatical phenomena related to the choice between the auxiliaries ESSERE ('be') and AVERE ('have') with non-transitive (unaccusative and unergative) verbs. It is argued that a purely syntactic account of unaccusativity is insufficient to capture the variation exhibited by these verbs. In particular, it is claimed that the unmarked selection of ESSERE with unaccusatives and of AVERE with unergatives in the present perfect tense is sensitive not only to a hierarchy of syntactic configurations (as assumed by the Government-Binding version of the Unaccusativity Hypothesis) but also to lexical hierarchies that subdivide the range of unaccusative and unergative verbs along semantic dimensions. Such hierarchies distinguish 'core', or prototypical, types of verbs from peripheral ones , and are consistent with the historical evolution of auxiliaries in Romance. However, auxiliary selection in syntactically marked 'restructuring' constructions, induced by certain Raising and Control verbs, is not sensitive to these semantic dimensions. It was predicted that the interaction between syntactic and semantic constraints would give rise to systematic variability in native speakers' linguistic intuitions, manifested in consistent and determinate acceptability judgments on core types of verbs, and variable or indeterminate judgments on peripheral types of verbs. It was also predicted that non-natives would differ from natives in terms of the extent to which indeterminate judgments penetrated from the periphery to the core.

Methodologically, this study represents the first application of magnitude estimation techniques to the elicitation of linguistic acceptability judgments. Magnitude estimation makes it possible to measure variability in acceptability judgments directly, which has the advantage of producing interval scales that can then be properly analysed by parametric statistics. Other ranking elicitation procedures produce only ordinal measurements. A systematic comparison between the judgments obtained by means of magnitude estimation and those obtained by means of a card-sorting ranking procedure indicates that both native and non-native speakers are able to judge acceptability via magnitude estimation with at least as much delicacy as via card-sorting. In some cases, magnitude estimation produces finer-grained distinctions of unacceptability, and reveals differences between native and near-natives which are not reproduced in the card-sorting task.

A series of experiments was conducted addressing the three issues of (a) variability in native intuitions, (b) progressive development of non-native knowledge, and (c) ultimate attainment at near-native competence levels. Acceptability judgments were collected from Italian native speakers and English-speaking learners of Italian at four proficiency levels (beginner, intermediate, advanced, near-native). A group of French near-native speakers of Italian was also tested for the purpose of comparison with the English near-natives.

The results show that (a) the judgments of native Italians are sensitive to different lexical-semantic hierarchies of unaccusative and unergative verbs: judgments on the basic syntactic reflexes of the unergative/unaccusative distinction (auxiliary selection and *ne*-cliticization) exhibit more or less determinacy depending on the semantic characterization of individual verbs; however, native speakers discriminate categorically between possible and impossible, obligatory and optional auxiliary change under restructuring, irrespective of the semantics of the inducing verb; (b) non-native judgments reflect a difference in learnability between lexical-semantic and purely syntactic distinctions. Lexical-semantic hierarchies affect the development and ultimate shape of non-native grammars, in that interlanguage representations for core lexical classes are constructed earlier than those for peripheral classes, with non-native acceptability values gradually approximating the native values. Peripheral restructuring constructions, however, never become determinate in the interlanguage grammars of English learners, which are incomplete in this respect even at the near-native level. In contrast, French near-native speakers of Italian show evidence of having constructed determinate, but divergent representations of these syntactic phenomena. It is argued that such differences in ultimate attainment reflect differences in the overall representations of unaccusativity in French and English.

LEXICAL CONDITIONS ON SYNTACTIC KNOWLEDGE: AUXILIARY SELECTION IN NATIVE AND NON-NATIVE GRAMMARS OF ITALIAN

PREFACE

This study grew out of my interest for two distinct objects of inquiry: linguistic intuitions, and Italian auxiliaries.

Knowing more than one language, I have often been intrigued by the elusive nature of linguistic intuitions, and particularly by the difference between clear and unclear intuitions. An adecdote can perhaps serve as an illustration of this distinction. A few years ago, when I was a student at USC, I received a telephone call one evening from a colleague who was working on Romance syntax. He was eager to finish a paper, and wanted to know my intuitions, as a native speaker of Italian, on a series of sentences that included particular syntactic constructions. It was urgent, he said, so I agreed to do this over the telephone. The task turned out to be more demanding than I had expected. While I did not have any doubts about some of these sentences, others required all my concentration, and even after hearing them three times I could not make up my mind. But my colleague was rather impatient, and insisted that all he wanted was a 'yes' or a 'no' despite my obvious indecision. Although I felt pushed to give him categorical answers, I felt that these did not adequately reflect my uncertainty. It seemed to me that there were interesting theoretical reasons to distinguish the unclear sentences from the clear ones. In treating all my judgments in a categorical way, my colleague had chosen to ignore those reasons.

As a reasonably fluent non-native speaker of English, I am often surprised at the number of grammatical phenomena which are obvious to natives, but about which I still do not have any clear intuitions. Clearly, years of exposure to English have equipped me with intuitions about a wide range of constructions, but have not been sufficient to provide a feeling for the acceptability of other properties of grammar. Is lack, or unclarity, of intuitions a permanent feature of non-native competence? Is my non-native indecisiveness different in nature from my native indecisiveness?

These anecdotal considerations point to two general problems with linguistic intuitions. First, the psychological nature of linguistic intuitions, and of the process that leads to specific acceptability judgments, is poorly understood. Second, there is an obvious discrepancy between the complexity of linguistic intuitions and the rudimentary methods with which they are treated as data in linguistics. Linguistic intuitions are not the focus of psychophysical exploration like sensory perceptions, or subject to psychometric criteria, like social psychological judgments. The elicitation of linguistic intuitions is not normally carried out on the basis of a quantitative model. The measurement scales onto which intuitions are

projected tend to be dichotomous (i.e. 'yes' or 'no') at worst, and ordinal (i.e. 'more' or 'less') at best: such scales cannot provide any theoretically interesting insights on whether differences in acceptability are large or small, or correspond to important or trivial factors in the mental representation of language; moreover, they do not readily permit the use of parametric statistics on the results.

One of the purposes of this study is allow this lack to be remedied, by approaching linguistic acceptability with a technique which has long been applied to psychophysical and social psychological judgments: magnitude estimation. This method is sensitive to variation in linguistic acceptability, because it makes it possible to measure acceptability directly on an interval scale.

Magnitude estimation could show its fruitfulness only if applied to an area of grammar in which variability of judgments is critical. This leads us to the other source of inspiration of this study: auxiliary choice with Italian intransitive verbs.

The systematic variation of auxiliary choice in Italian has been explored in detail in the linguistic literature. Italians say either 'ha piovuto' (it has rained) or 'è piovuto' in free variation, but they know that only 'mi è piovuto sulla testa' (it has rained on my head) is possible; they say 'ha passeggiato' (she has walked), but 'è riuscito' (she has succeeded). I tend to say 'Maria è voluta andare all'università' (Maria has wanted to go to the university) but many Italians, and most non-native speakers of Italian, say 'Maria ha voluto andare all'università'. At the same time, I have no doubts that 'E' andare all'università che Maria è voluta' (it is to go to the university that Maria has wanted) is wrong and 'E' andare all'università cha Maria ha voluto' is right, whereas many English friends with excellent Italian do not know the difference. These and other facts have been accounted for by syntactic and lexical-semantic theories of auxiliary selection, which have offered revealing explanations of the general mechanisms that govern auxiliary choice. The single major insight underlying most theoretical argumentation - the Unaccusative Hypothesis - elegantly captures major differences with respect to auxiliary selection and other syntactic properties by setting up a distinction between two sub-classes of intransitive verbs: unaccusative and unergative.

However, there is another kind of variation in auxiliary behaviour that is usually neglected, or dismissed as unimportant: variation WITHIN the classes of unaccusative and unergative verbs. This study proposes that decomposing the classes of unaccusative and unergative verbs into semantically coherent sub-types reveals a common pattern underlying both diachronic and synchronic variability in auxiliary selection. It is shown, for example, that the

well-attested process of historical change in Romance auxiliaries has been expanding the domain of the 'have' auxiliary (at the expense of the 'be' auxiliary) spreading from the periphery towards the core of a hierarchy of unaccusative verb types. This hierarchy determines predictable variation in native speakers' linguistic intuitions on unaccusative verbs, determines a directionality of difficulty in the acquisition of ETRE by native Italian speakers compared to the acquisition of ESSERE by native French speakers, and shapes the development of the ESSERE auxiliary in second language grammars of Italian.

In short, then, this thesis is an experimental approach to systematic variation in native and non-native intuitions on Italian auxiliary selection. It is my hope that it provides a theoretical contribution to knowledge of language and of second language acquisition, and that it encourages systematic research on the nature of linguistic intuitions.

CHAPTER 1
INDETERMINACY IN LINGUISTIC INTUITIONS

1.1 Native intuitions and linguistic theory

> "In actual practice, linguistics as a discipline is characterized by attention to certain kinds of evidence that are, for the moment, readily accessible and informative: largely, the judgments of native speakers. Each such judgment is, in fact, the result of an experiment, one that is poorly designed but rich in the evidence it provides. In practice, we tend to operate on the assumption, or pretense, that these informant judgments give us 'direct evidence' as to the structure of the I-language but of course this is only a tentative and inexact working hypothesis...in principle, evidence concerning the character of the I-language and initial state could come from many different sources apart from judgments concerning the form and meaning of expressions: perceptual experiments, the study of acquisition and deficit or of partially invented languages such as creoles, or of literary usage, or language change, neurology, biochemistry and so on...(but) the judgments of native speakers will always provide relevant evidence for the study of language...although one would hope that such evidence will eventually lose its uniquely privileged status."
> (Chomsky 1986: 36-37)

While it is still a widespread practice on the part of linguists to rely on grammaticality judgments in order to support their theoretical claims, there has been a growing awareness of the fact that very little is known about the psychological nature of linguistic intuitions. The aim of this chapter is to put forward the case that a clearer understanding of the cognitive factors involved in the internal origin of linguistic intuitions and in their overt expression as judgments is essential for a more effective and informative exploitation of intuitional data.

Criticisms of acceptability judgments as linguistic evidence have been mainly concerned with two issues. The first - and perhaps most fundamental - is the **validity** of judgments, or the relationship between (i) linguistic intuitions and grammatical competence, (ii) linguistic intuitions and acceptability judgments, and (iii) linguistic intuitions and different kinds of underlying norm.

The second issue is the **reliability** of judgments, referring to (i) consistency among native speaker judgments and to (ii) the indeterminate grammaticality of certain areas of language. The concept of *indeterminacy* is one of the cornerstones of this thesis. Provisionally, let us define it as the indefinite grammaticality status of certain constructions in the speaker's

internalized grammar, which leads to variable and/or indecisive acceptability judgments. These issues will be discussed at length later in the chapter.

1.1.1 The question of validity

If the elicitation of intuitional data is regarded as a small-scale experiment, the results of which are individual judgments, then it is crucial to ensure that such judgments actually tap the speaker's internalized grammatical competence. The assumption that there exists a correspondence between judgments and underlying grammar has been the target of several objections.

1.1.1.1 Grammatical competence and extralinguistic factors

Two arguments have been raised against the supposed validity of linguistic intuitions as indicators of competence:

(1) the capacity to have relevant intuitions may not be a reflection of grammatical competence: it may derive from a separate faculty characterized by a set of properties *sui generis* that are not shared by other kinds of linguistic behaviour (Bever 1970, 1974; Snow and Mejer 1977; Gleitman and Gleitman 1979)[1].

(2) even if linguistic intuitions are directly related to grammatical competence, they may be affected by other factors that are extralinguistic in nature and cannot be easily isolated. A sentence may be judged acceptable or unacceptable for reasons that have little to do with its status in the competence of speakers - in other words, speakers may direct their attention towards aspects of the sentence irrelevant to the purpose of the experiment and judge something different from what they were expected to judge (Botha 1970, 1973; Levelt 1974).

The former argument is more fundamental than the latter, since if linguistic intuitions turned out to be totally (or even largely, if unpredictably) independent of the speaker's internalized grammar it would obviously make no sense to use them as basic data for the purpose of constructing models of grammar. Although the psychological laws of the intuitional process are poorly understood, it is indisputable that the use of acceptability judgments and introspective reports has led to the establishment of a substantial number of significant generalizations about syntactic processes (see Newmeyer 1983 on this point). These results would hardly be explainable if no more than a chance relationship was assumed between

grammatical knowledge and expressed linguistic intuitions. Moreover, acceptability judgments and linguistic performance have often been shown to be highly correlated (Quirk and Svartvik 1966; Greenbaum and Quirk 1970): this suggests that native speakers tend to rely on the same grammar for both the sentences that they accept and those that they are able to produce.[2] There are therefore sufficient grounds to disregard the claim that there is no orderly relationship between linguistic competence and intuitional processes, and between intuitional processes and performance.

The problem still remains that the speaker's internalized grammar is not the only system activated in the intuitional process.The interaction of the grammar with other cognitive and pragmatic systems is fully explainable within a modular conception of language, according to which the grammar is only one of a number of human cognitive systems, each governed by its own principles and each contributing to the superficial complexity of language (Newmeyer 1983; White 1982). Botha (1983) reserves the term 'spurious' for intuitions determined or affected by extragrammatical factors, as opposed to 'genuine' intuitions, which originate from the informant's internalized grammar. A variety of factors may be at the source of spurious intuitions. To mention the most relevant:

(a) perceptual strategies (Bever 1970, 1974; Snow 1974), as in the famous garden path example "The horse raced past the barn fell", which tends to be judged as ungrammatical by most informants because of the tendency to take the first NP V NP sequence as the main clause. Difficulty of parsing may be responsible for the rejection of perfectly grammatical sentences; conversely, ease of parsing may lead to the acceptance of sentences that were predicted to be ungrammatical on theoretical grounds;

(b) context of presentation (Levelt 1971, 1974; Snow 1974; Greenbaum 1977). A sentence of dubious grammaticality is more likely to be judged as ungrammatical if placed after a set of clearly grammatical sentences, or as grammatical if following a set of clearly ungrammatical sentences.[3] This suggests that judgments in isolation are very different from judgments by contrast, as will be pointed out in Chapter 2; furthermore, the order in which sentences are presented to informants may influence their judgments;

(c) pragmatic considerations: when faced with syntactic ambiguity, informants tend to prefer the reading that (1) represents the most frequent interpretation, and (2) requires fewer assumptions about previous discourse (Hawkins 1982; Altmann and Steedman 1988). The importance of context for certain kinds of constructions is often underestimated.

Some decontextualized sentences may be judged in arbitrarily different ways, depending on the context provided by the informant in order to interpret them.[4]

(d) mental, or introspective state. Speakers have been found to have different intuitions on the same sentences, depending on whether they were facing a mirror. The theory here distinguished subjective and objective self-awareness (Carroll, Bever and Pollock 1981). These states were experimentally manipulated by the presence or absence of a mirror in front of the informants. Objective self-awareness induced attention to the semantic aspects of sentences, whereas subjective self-awareness was conducive to attention of syntactic properties.[5]

(e) linguistic training (Levelt 1974; Botha 1973). When linguists use themselves as main (or only) informants (producing "the theory and the data at the same time", as Labov (1972) puts it) there cannot be any guarantee that their judgments are not biased by theoretical expectations. Also, linguistically naive native speakers have been found to be different in their acceptability judgments from linguistically sophisticated native speakers (Spencer 1973; Snow and Meijer 1977; Gleitman and Gleitman 1979), although it is difficult to pinpoint the exact nature of such difference: naive informants appear to be more normative and more confident (Ross 1979), or less consistent (Snow and Mejer 1977; Bradac et al. 1980).

The problem that linguists working with intuitional data have to face is therefore the following: linguistic intuitions do reflect grammatical competence but they are also open to the influence of other cognitive or contextual factors. The extragrammatical factors can be controlled for, at least to a certain extent, by carefully selecting the test sentences, the test design, and the informants. Even if the tightest precautions are adopted, however, one can never be absolutely sure that intuitions are derived exclusively from grammatical competence.

1.1.1.2 'Acceptability' vs 'grammaticality'

The terms 'acceptability' and 'grammaticality', although often confused, correspond to two distinct points of view on the status of sentences. From the point of view of the linguist, sentences may or may not be grammatical with respect to a particular formal representation of competence (ie. a linguistic theory). The term 'grammaticality', however, is inappropriate to describe to the feelings (naive) informants have about the well-formedness of sentences: to them sentences are only acceptable with respect to the various variables (grammatical

competence, metalinguistic knowledge, pragmatic appropriateness, etc) that, as we have just seen, may determine their judgments. Thus, if speakers are asked to judge whether a sentence is grammatical without any indication of what grammaticality means, they will attribute different meanings to the term (interpretable, contextualizable, possible in a given dialect, correct with respect to prescribed rules, etc). There often is a conflict between grammaticality, as predicted by linguistic theory, and acceptability: sentences that are grammatical may be judged as unacceptable, and vice versa (Cowart 1989).[6]

1.1.1.3 Underlying rules and normative reactions

Even if it is assumed, for the sake of simplicity, that all extralinguistic factors can be isolated, a decision about the nature of the rules underlying genuine intuitions is not a straightforward task. In producing acceptability judgments speakers may unconsciously shift towards the norm they believe they should follow, and away from the norm actually governing their internalized grammar (Greenbaum and Quirk 1970; Coppieters 1987). It is important to distinguish among different *attitudes to usage* that speakers may have. According to Greenbaum and Quirk, three potentially different but often interacting factors can be reflected by speakers' judgments: (a) beliefs about the forms they habitually use; (b) beliefs about the forms that ought to be used, and (c) willingness to tolerate usage in others that corresponds neither to their own habitual forms nor to prescriptive forms.

Speakers' conscious beliefs about language may lead them to formulate adaptive rules that, under certain circumstances, modify the output of their mentally represented grammars, often in order to avoid the production of stigmatized forms (Anderson 1973; Pateman 1985). Adaptive rules are cognitively different from 'tacit' rules belonging to the internalized competence: they are social or cultural in character and usually more accessible to introspection than internalized rules.[7] Adaptive rules may therefore be more readily available to informant as a basis for acceptability judgments.

1.1.1.4 Intuitions and judgments

It may also be appropriate to distinguish between *intuitions* and *judgments* , although the terms are often used interchangeably. Linguistic intuitions ("non-reasoned feelings": Botha 1973) are not easily accessible, either to the speaker's conscious mind or to the researcher: they can be portrayed as the result of a computational process that is said to take place in the speaker's mentally represented grammar, and below the level of conscious awareness. Intuitions are registered and reported by the speaker in the form of judgments, which are

linguistic *descriptions* and may therefore be inaccurate. This applies in particular to judgment tests requiring complex verbalizations of rules or other forms of metalinguistic statements.

1.1.1.5 Correctness of intuitions

A different but equally basic validity question has been raised with respect to the possibility of assessing the correctness of intuitions (Ringen 1975). When requested to judge the acceptability of a given sentence, even those speakers who understand the task in the intended manner and produce their judgments on the basis of their grammatical competence may nevertheless make errors. The researcher has no principled means of checking the correctness of intuitions in this particular sense. The implicit practice is to assume that informants have *final epistemic authority* on their intuitions, in the same way as they have final epistemic authority on their sensations (which cannot be disproved, since mental states like sensations are in principle uncheckable). Ringen suggests, however, that if judgments are to be regarded as empirical facts, then it should be possible to verify their correctness: this would imply measuring the disparity between the informants' judgments and the objective physical states of affairs about which judgments are made, or, in other words, having at one's disposal an assessment procedure independent of the reports of the speakers whose judgments are being assessed. For the correctness of judgments to be empirically assessable, it should be possible to measure intuitions of degree of grammaticality against some independently established grammaticality scale. But since intuitions are used as primary data for the construction of models of grammar, obviously there cannot be such a grammaticality scale independent of intuitions.

Thus, in order to measure linguistic acceptability, one has to decide what kind of dimension acceptability is. Important implications follow from this initial decision. Although acceptability is obviously not a physical dimension like length, brightness or weight, it cannot be automatically equated with pure 'sentiments' or sensations either.[8] To the extent that acceptability is a function of grammatical competence, and on the assumption that grammatical competence derives from a mental faculty characterized by physical and biological properties, linguistic perceptions should be subject to the same psychophysical laws as other perceptions: it would then be possible to measure the disparity between acceptability judgments on given aspects of the grammar and the objective representation of those aspects in the mind. In practice, however, there are no ways of measuring the mental representation of linguistic knowledge other than asking informants to provide acceptability judgments. We will come back to the empirical considerations involved in measuring linguistic acceptability in Chapter 2.

1.1.2 The question of reliability

Reliability is usually related to consistency: both inter-subject consistency, that is, agreement among judgments produced by different informants, and intra-subject consistency, or agreement among judgments produced by the same informant in different replications of the test.

1.1.2.1 Inter-subject consistency

It is well-known that there are areas of grammar on the acceptability of which native speakers do not agree, or do not have any clear intuitions. The easiest - and least satisfactory - solution to the question of inter-subject inconsistency is to ascribe conflicting intuitions (assuming that they are genuine) to idiolectal or dialectal differences. In other words, speakers may disagree in their intuitions because they do not share the same grammar. According to this position, there cannot exist any in-built variation in the grammar. Ringen (1974) distinguishes between the non-mentalist view of language, according to which language is a cultural object or an institution, and the mentalist view, which regards language as the mental reality underlying speech behaviour. Interestingly, *both* views downplay the importance of inter-subject consistency of judgments as a criterion in linguistic research, either because members of the same speech community are supposed to share the same cultural institutions (the 'logic of Saussure" in Labov's (1972) definition) or because speakers of the same language are supposed to share the same mentally internalized grammar.[9]

Another common solution is to minimize the inconsistencies of native speakers' intuitions. Newmeyer (1983) draws a distinction between *superficial* and *genuine* disagreement about sentence grammaticality. Superficial disagreement does not concern the actual grammaticality status of a given sentence but rather its analysis: whether, for example, some fundamental characteristics of a sentence should be explained by a grammatical principle or by an extra-grammatical one. Genuine disagreement, on the other hand, arises from conflicting intuitions about the grammaticality of the sentence. In Newmeyer's view, the vast majority of alleged native speakers' disagreements on data are "essentially theoretical disagreements that are only superficial disagreements about the acceptability of sentences." (Newmeyer 1983: 66).

1.1.2.2 Intermediate grammaticality

Cases of genuine disagreement, however, are far from uncommon. Early theories of transformational grammar implicitly recognized the existence - if not the extent - of this problem, although their solution was based on the concept of definite grammaticality ("...in many intermediate cases we shall be prepared to let the grammar itself decide, when a grammar is set up in the simplest way so that it includes the clear cases and excludes the clear non-cases." (Chomsky 1957: 14)).[10]

The alternative way of accounting for inter-subject disagreement is to assume that native language grammars are indeterminate and that such indeterminacy is a characteristic of natural human languages. In this perspective, linguistic structures are not simply grammatical or ungrammatical: they may be grammatical to a degree. From the point of view of the informant, structures may have varying degrees of acceptability, which may determine different degrees of consistency in her judgments. Conceptually, indeterminacy may be defined as 'indefiniteness of status in the speaker's grammatical competence'. Operationally, indeterminacy may be defined as 'variability in the speaker's acceptability judgments'. Such variability may manifest itself in inconsistency of judgments (both within- and between speakers), or in an inability to distinguish acceptable from unacceptable sentences. The question that concerns us here is whether indeterminacy is random and unpredictable, or whether it is lawful and predictable.

1.1.2.3 'Fuzzy' grammars and acceptability hierarchies

Outside the generativist framework, models of grammars have been elaborated on the basis of grammaticality as a *relative* property of sentences. Lakoff (1973) and Mohan (1977) maintain that there exists an ordinal scale of acceptability within a speech community such that all speakers are likely to agree on the rank order of acceptability values in a given set of sentences, although they may disagree on the absolute rating of individual sentences. The pattern of individual rating judgments, therefore, should not be random but should reveal an implicational scale, the basis of which is a shared ordinal scale. In other words, it would be inconsistent for a speaker to rank sentence A as more acceptable than sentence B but then accept B and reject A.[11]

Along the same lines, Ross (1979) suggests that a language L may be seen as consisting of an indefinite number of acceptability hierarchies, each leading away from the core to the periphery of L, and governed by implicational laws such that (a) acceptance of a sentence at

distance x from the core implies acceptance of any more central sentences along the same hierarchy (i.e. between x and the core), and (b) speakers may disagree on the absolute acceptability values of individual sentences because they may have different acceptability thresholds on the same hierarchy.[12]

1.1.2.4 Acceptability hierarchies and parametric variation

The concept of acceptability hierarchy relies on the distinction between a core and a periphery . While early generative grammar could not admit this distinction without violating some of its basic premises, recent developments within a principles-and-parameters framework do make frequent reference to it, thus allowing for the existence of language indeterminacy - at least in principle - in a more satisfactory way. However, the research to date has done very little to substantiate the notion of periphery, or to investigate the modes of its interaction with the core. Furthermore, statements concerned with the periphery seem to contain a fundamental ambiguity with respect to the nature of the relationship between core and periphery: at times, this is seen as a *discontinuous* relationship, whereas at other times it is regarded as a *continuous* relationship.

It is recognized, for example, that the periphery may contain "borrowings, historical residues, inventions and so on" (Chomsky 1981: 8), and that

> "we would expect phenomena belonging to the periphery to be supported by specific evidence of sufficient 'density', to be variable among languages and dialects, and so forth... What we 'know innately' are the principles of the various subsystems of [the initial state] S_0 and the manner of their interaction, and the parameters associated with these principles. What we learn are the values of the parameters and the elements of the periphery (along with the lexicon, to which similar considerations apply)."
> (Chomsky 1986: 147, 150).

These statements suggest that acquiring the periphery of any language may require substantially different evidence from that necessary to acquire the core, and thus they implicitly subscribe to the discontinuity assumption.

Elsewhere, however, it is said that "the distinction between core and periphery leaves us with three notions of markedness: core versus periphery, internal to the core, and internal to the periphery." (Chomsky 1986: 147). This statement, unlike the previous ones, contains implicit support for the continuity assumption. If the periphery is contiguous with the core,

then it is possible to see both as the extremes of the same scale of markedness. Fodor (1989) argues that this view has at least four advantages. In her words,

> "(i) On the assumption of continuity, facts about the periphery can provide valuable constraints on our theories of the core. If two or more alternative formulations of the principles appear to be equally compatible with the facts of the core, we can select between them on the basis of which most appropriately ranks other phenomena as more and less peripheral.
>
> (ii) The continuity assumption offers fewer degrees of freedom, with the result that success in the characterization of the class of human languages would be more convincing, and even failure would be more informative. Specifically, continuity renounces a wide range of otherwise available choices - choices about where the dividing line between core and periphery falls, and about how the systems on either side of the line differ from each other.
>
> (iii) There appears to be no shared pretheoretical hunch that a sharp dividing line exists (*N.B.* within 'pure' syntax; the lexicon does present a fairly clear distinction between generalizations and exceptions).
>
> (iv) If the continuity assumption could be upheld, the parsimony of the parameter-setting model of acquisition would extend from the core to the periphery. A discontinuous periphery would require, in addition, all the traditional psychological mechanisms for hypothesis formation and testing."
> (Fodor 1989: 131)[13]

The continuity assumption has another fortunate consequence: it allows us to say that indeterminacy is naturally compatible with Universal Grammar (UG), and offers an explanation for acceptability hierarchies. With Pateman (1985), we can distinguish among
(a) cases in which UG rigidly specifies the form of grammar independently of any evidence in the input. Principles such as structure-dependency are well known to constrain linguistic development in such a way that, regardless of the amount and type of evidence they are exposed to, children never make errors that contain violations of these principles;
(b) cases in which UG offers a preferred structure for an array of input data; the acquisition of such marked structures requires positive evidence. This claim does not imply that alternative solutions may not be entertained by the learners.For example, although there is a tendency

towards regularization, irregular forms are eventually incorporated in the grammar and passed on to the next generation of speakers;
(c) cases in which UG offers a small, unordered set of structures as analyses for given input patterns, presumably because of some (still largely unknown) relaxation of UG constraints (Chomsky 1981). This latter type of structures properly belongs to the periphery, and allows for unpredictable elements of variation to enter linguistic reproduction (see 1.1.2.2.3).
Thus, we can say that properties closer to the core extreme of an acceptability hierarchy occupy that position because they are specified by UG, without possible alternatives; properties that fall in the middle are also determined by UG but options are possible; peripheral properties are indeterminate, since they are not fully, or not at all, specified by UG. It is not implausible that UG and cognitive resources non-specific to language have a combined influence on the periphery, although the modes of their combination are still largely unexplored. This means that cultural norms, speech adaptations, individual beliefs and conscious rationalizations about language would find a more fertile ground at the periphery than at the core, and accounts for the fact that the maximum amount of variation and inconsistency is usually found in this area.

1.1.2.5 Indeterminacy in a diachronic perspective

The concept of indeterminacy is also relevant to the question of language change. Regardless of whether the introduction of innovative change happens abruptly and discontinuously with respect to previous stages of the evolution of a language (see Lightfoot 1983, 1990)[14], or whether it implies a gradual and cumulative process (Traugott and Heine 1991),

> "....there will always have to be at least one initiating change motivated externally in terms of changes in the community that provides the data to which the learners is exposed...These changes may be determined either by variation in adults, or at least postmaturational usage, or by population shifts and other kinds of socially motivated factors."
> (Vincent 1989, in Lightfoot 1989)

Consistently with this view, language change can be regarded as a process that starts by creating indeterminacy at the periphery of acceptability hierarchies, which then gradually extends towards the core. Such indeterminacy steadily increases, affecting the frequency and consistency of appearance of certain forms in the input, and therefore the triggering experience of children in the course of language acquisition. This process would reach a point when, in the course of generations, abrupt reanalysis of learners grammars becomes

necessary because the affected forms are in disharmony with other structural properties of the language, and a grammar containing such forms is unlearnable (Lightfoot 1990; Kroch 1989 in Lightfoot 1989).[15] We shall return to these issues later.

1.1.2.6 Psychometric plausibility of acceptability hierarchies

Finally, the concepts of relative grammaticality and acceptability hierarchy are important because, besides accounting for intersubject disagreement in judgments and for lack of intuitions, these notions are also consistent with the fact that people are usually better at producing *relative* rather than *absolute* judgments (see Nunnally 1976: 44 and Chapter 2 of this thesis). This seems to be true of acceptability judgments, even when they involve rating isolated sentences. When a given form has close variants, informants may match the variants mentally before making a judgment (Greenbaum 1988), or try to find a visual context in which the sentence could make sense (see Levelt et al. 1977, who suggest that acceptability judgments tend to be faster and more positive for high imagery, or concrete materials than for low imagery, or abstract materials). If so, then inconsistencies in absolute acceptability ratings may be due to differing abilities of individuals to retrieve a matching variant, or to construct mental contexts.

1.2 Non-native linguistic intuitions

The single most important notion underlying the field of second language acquisition research is that of interlanguage. Clearly inspired by generative linguistics, the 'Interlanguage Hypothesis' (whose original formulations are found in Corder 1974, 1981; Selinker 1972) assumes that second language learners have a mentally represented grammar at every stage of the acquisition process. This grammar is different from both the native language and the target language (hence the term 'interlanguage'). Interlanguage grammars can be studied in their own right (as any fully developed natural language) and not necessarily in terms of a comparison with the target language. The state of learners' interlanguage competence can be formally characterized by a grammar that can (minimally) meet the level of descriptive adequacy, when it correctly accounts for the data produced by learners, or the higher level of explanatory adequacy, when it explains how the learners came to acquire that competence. Most early research on second language acquisition was devoted to the former aim (for example Dulay and Burt 1973; Ravem 1974; Cazden, Cancino, Rosansky and Schumann 1975; Wode 1978), whereas recent theories developed within the principles-and-parameters framework address the latter (see White 1989, and the

collections in Flynn and O'Neill 1988; Pankhurst, van Buren and Sharwood Smith 1988; Eubank 1991, White (forthcoming), among others).

If learners are assumed to have interlanguage internalized grammars, then learners' linguistic intuitions become the primary indicators of interlanguage competence. The problem then arises of whether it is appropriate to treat native and non-native competences as identical objects of investigation. It is easy to get the impression, in reading the second language acquisition literature, that principles and methodologies have been borrowed from linguistic theory without questioning their validity and their applicability to the study of interlanguage grammars. Given the fundamental difference between native and non-native grammars - that the former are, at least in their core, fully developed, steady states of linguistic competence, whereas the latter are unstable, transitional states of knowledge, it is even more surprising that very few studies (among which Birdsong 1989; Chaudron 1983; Sorace 1988) have so far attempted a more precise definition of non-native linguistic intuitions, or of the elicitation procedures employed in the collection of intuitional data from second language learners.

1.2.1 Validity and reliability of non-native intuitions

Learners' acceptability judgments are generally assumed to be related both to (a) the transitional state of their interlanguage knowledge, and to (b) their actual intuitions. The same arguments underlying the question of the validity of native acceptability judgments apply to non-native judgments. Thus, it seems reasonable to claim that interlanguage judgments at least partly reflect interlanguage knowledge: the more extraneous variables are controlled for, the clearer this reflection is (see the arguments outlined in 1.1.1.1 above).
It can be a more complex task, however, to decide about the kind of *norm* consulted by learners in the process of producing a judgment, particularly in a learning environment that fosters the development of metalinguistic knowledge: it is difficult to tell whether subjects reveal what they think or what they think they should think. The elicitation of immediate judgment responses under well-defined time constraints may provide a partial solution to this problem, as we will show later.
Moreover, the verbal articulation of linguistic intuitions can be a more demanding task for learners than it is for native speakers. It is therefore advisable to restrict judgmental tests to simple recognition tasks (Kohn 1979, 1982), at least with learners at low levels of proficiency.

A further difference between native and non-native intuitions concerns the correctness issue. We saw earlier that native acceptability judgments can be simultaneously regarded as proper judgments and as sentiments: on the one hand, they result from grammatical

competence, but on the other hand they cannot be compared to objective states of grammatical competence because competence has so far eluded direct measurement. Non-native acceptability judgments can, at first glance, be compared with the objective reality represented by native judgments; this is indeed common practice in most experimental studies in second language research, where a control group of native speakers serves as the basis for comparison. However, there are two considerations to bear in mind. First, native judgments themselves can be indeterminate, particularly when the objects of investigation are highly marked or very subtle syntactic properties: native judgments may therefore provide a reference point only in the most uncontroversial cases. Second, the ultimate purpose of testing non-native linguistic intuitions is to construct a model of the non-native grammar of a particular learner, or group of learners at a particular level of proficiency : unless such non-native grammar is seen as an imperfect realization of the native grammar, the learners' judgments themselves should provide the primary criterion for deciding which structures are or are not part of it. The evaluation of the distance between native and non-native grammars becomes an irrelevant criterion.

1.2.2 Interlanguage indeterminacy

What characterizes interlanguage grammars and distinguishes them from native grammars is the pervasiveness of indeterminacy. In many cases, the indefiniteness of interlanguage competence leads to the learner's inability to express a clear-cut judgment of acceptability.

At the most basic level, constructions are indeterminate because the learners do not have any knowledge of them.[16] This kind of indeterminacy characterizes non-native grammars throughout the acquisition process, although it is more conspicuous at the initial and intermediate stages of interlanguage development.

At more advanced stages, constructions may become indeterminate (after a period of relative stability) because of the increased amount and sophistication of the learner's knowledge. One of the factors contributing to this kind of interlanguage indeterminacy is the permeability of interlanguage grammars, their 'openness' to the penetration of other linguistic systems. It has been debated whether permeability is a competence phenomenon, or it is restricted to the performance level (Adjémian 1976, 1982; Liceras 1983); or whether permeability is peculiar to interlanguage but does not affect native language grammars (Arditty and Perdue 1982). If permeability is a crucial, though possibly not unique, property of interlanguage *competence* then it generates indeterminacy by creating the conditions for the coexistence of more than one rule for the same aspect of grammar. The coexisting rules may belong to

different linguistic systems known to the learner, or to successive stages of interlanguage development, or both. The result is variability or indecisiveness in learners' intuitions.

As Klein (1986) suggests, interlanguage grammars may be regarded as 'test grammars' in which a rule is associated, at a given time T, with both a degree of confirmation, indicating the confidence with which the learner knows a particular structure, and a degree of criticalness, representing the stability of that rule: in other words, whether that rule is undergoing a process of change at that time. The relation between confirmation and criticalness is not necessarily linear although, as Klein point out, weakly confirmed rules are more likely to become critical. Naturally, the criticalness of rules may increase only if interlanguage grammars are permeable to new input. The ever-changing nature of interlanguages necessarily brings about diachronic variability, and therefore makes rules critical. Rules cease to be critical when interlanguage grammars are no longer permeable, i.e. when a rule has been acquired or when a rule has become fossilized.

1.2.2.1 Conflicting parameter-settings / conflicting cognitive faculties

From a UG perspective, a further cause of indeterminacy can be identified in the fact that the UG-driven specification of core properties is narrower in scope and strength than in native grammars. This under-specification may be due to the reduced availability of UG as a cognitive module, or to inadequate exposure to input, or - in most cases - to both.[17] The result is a wider periphery and consequently more room for permeability and variation. Let us examine the latter point in more detail.

While UG, as a language-specific cognitive module, constrains first language acquisition, finding uncharted territory, in the presence of rich input, for fixing the parameter settings relevant to the language being acquired, the typical second language acquisition situation, allows either (or more often both) of the following conditions to obtain:

(a) there may be potential attrition between the parameter settings of the learner's native language and those required by the L2. Depending on the nature of the evidence available and on the specific markedness relationship between the native and the target language,[18] L2 parameters may have to be reset, or simply left unset (see White 1988, 1989, 1990; Flynn 1987; Schachter 1988 on these issues). The quantity and/or quality of the input available may be insufficient for the L2 parameters to be (re)set and for significant projections to take place.

(b) outside the scope of UG, there is competition between UG and problem-solving cognitive resources generally available to adult learners (Felix 1985), which may suggest

solutions to input configurations not definable in terms of UG constraints (often called 'unnatural' or 'wild' rules).

This conflict may account for the fact that acceptability hierarchies for a second language never become fully determinate: only a reduced portion of the L2 core is unambiguously specified by UG but a relatively larger portion has been (or continues to be) the battleground of conflicting parameter settings, and presents long-lasting or even permanent parametric variation. Furthermore, an ample periphery remains open to the influence of different kinds of factors, both grammatical and extra-grammatical in nature. If so, this would partly explain why variable or inconsistent intuitions are found in non-native speakers even at very advanced stages of language proficiency. To the extent that interlanguages are natural languages, falling within the range of possible grammars allowed by UG, learners intuitions will vary around a limited number of alternatives ('possible' alternatives); to the extent that interlanguages are also determined by problem-solving strategies, learners' intuitions will take more idiosyncratic and largely unpredictable forms.

Where the input presents severe limitations in terms of quantity and variety, as in conventional instructional settings, the scope of action of the language-specific faculty is further restricted by the lack of triggering evidence. Consequently, alternative structural solutions are likely to be provided either by the L1 or by cognitive resources not specific to language. These solutions in turn determine (a) greater inter-subject variability, since different learners may come up with idiosyncratic hypotheses, and (b) greater intra-subject variation, since individual learners may formulate competing hypotheses that cannot be solved by positive confirmation or disconfirmation, as would happen in naturalistic acquisition.[19]

1.2.2.2 Near-native indeterminacy

Does interlanguage indeterminacy invariably decrease over time, and eventually disappear? This question is intimately related to the issue of completeness in second language acquisition. One of the most controversial issues in second language acquisition theory is whether it is possible for the adult learner to construct native-like mental representations of grammatical knowledge for core aspects of the target language - or, in other words, whether non-primary language acquisition can be 'complete'. This would entail that non-native speakers construct acceptability hierarchies that are indistinguishable from those shared by native speakers.

This question is closely related to the ongoing debate on the availability of Universal Grammar as a constraint on interlanguage and the learnability of specific grammatical aspects of a second language: if interlanguage development was constrained by UG in its entirety, there would be no cognitive obstacle to complete success in the second language acquisition enterprise. The existing empirical findings, however, do not allow unequivocal interpretations. On the one hand, they show that interlanguage grammars are underdetermined by the input, since L2 learners can and do acquire subtle and complex linguistic knowledge that goes beyond the evidence available to them.[20] On the other hand, they provide indisputable evidence suggesting that many L2 learners fail to acquire grammatical properties that are not instantiated in their L1, or fail to re-set parameters after the initial, inappropriate adoption of the L1 setting (see White 1989 for a comprehensive account of the research to date). While both Plato's and Orwell's problems in second language acquisition are still waiting for a satisfactory explanation,[21] experimental evidence and intuitive observation at the present stage seem to be consistent in suggesting that - whatever the reason - reaching native-like competence with respect to *the whole* of the L2 grammar is an impossibility for the adult learner.

Most studies, however, do not address the question of ultimate attainment directly, since they are concerned with learners at the intermediate or advanced level . Surprisingly few studies look at the syntactic competence of near-native speakers - that is, learners at the most advanced stage of interlanguage development. Yet the 'ultimate attainment' issue can only be properly addressed by looking at the competence of learners who have supposedly reached that stage. One of the few exceptions is Coppieters's (1987) study, which aimed at investigating competence differences between the intuitions of native speakers and those of near-native speakers of French about some highly productive aspects of grammar, such as the contrast imperfect/present perfect, the distinction between the third person pronouns *il/elle* and *ce*, and that between preposed and postposed uses of adjectives. The most striking result of this study was the extent of the gap between native and near-native intuitions: while native speakers seemed to share a 'native majority norm', with respect to which they showed minimal variation, all near-native speakers deviate from the native norm in statistically significant ways (the closest non-native value to the majority norm was three stardand deviations from it), and revealed extensive variation. Furthemore, the *interpretation* that near-native speakers produced of the grammatical forms in question was on several occasions remarkably different from the interpretation offered by native speakers, even when their actual rating judgments were the same. This suggests, according to Coppieters, that the two groups of informants may have developed "significantly different grammatical systems for French", despite the fact that they are virtually indistinguishable in production. Finally, near-

native speakers often lacked any clear intuitions on some of the grammatical rules investigated (in particular, the distinction perfect/imperfect) and their preference for either form was unsystematic.[22]

The immediate conclusion that one could draw from this study is that similarity of performance between natives and near-natives may be a misleading indicator of near-nativeness at the competence level. Less obviously, the variety of near-native knowledge patterns revealed by these results suggests that the concept of near-nativeness may be more complex than it appears.The sense in which 'near-nativeness' is usually understood is that of *incompleteness*, corresponding to the absence in the near-native grammar of a property required by the native grammar. The term itself "*near*-native"seems to contain an implicit statement on the question of ultimate attainment, since it implies the non-coincidence of native and non-native grammars: the near-native grammar is *almost the same* as the native grammar, but falls short of it.

There is, however, another possible meaning of 'near-nativeness': the near-native grammar could be not so much incomplete as *divergent*, containing representations of L2 properties that are *different* from the native representations. One might expect to find a distiction between incompleteness and divergence manifested in the near-native speaker's grammaticality judgments. The incomplete grammar, lacking a given L2 property P, would lead to random, inconsistent, in short *indeterminate* judgments about P, whereas the divergent grammar, since it incorporates an alternative representation of P, would lead to *determinate* judgments that are consistently different from native judgments. Both the incomplete and the divergent near-native grammars could produce seemingly native-like performance, particularly if performance is based on a restricted range of structures that does not include P. But the undifferentiated use of the term 'near-native' may conceal two distinct states of grammatical competence, corresponding to two qualitatively different kinds of ultimate attainment.

Finally, it is worth noticing that near-native grammars may also be indeterminate in the same sense as native grammars, i.e. because of the intermediate grammaticality of certain constructions. This is the natural consequence of near-native speakers' success in acquiring acceptability hierarchies that are similar to those of native speakers. Given that, as we saw earlier, this type of indeterminacy leads to inconsistent and variable judgments, it may empirically difficult to decide whether inconsistency in near-native judgments is due to incompleteness or to intermediate grammaticality.

1.3 Conclusions

This chapter illustrated some of the conceptual issues underlying the use of acceptability judgments in linguistic research. It was argued that natural language grammars are characterized by indeterminacy, which corresponds to the unclarity of the grammaticality status of certain constructions in the speaker's internalized competence, and can be operationally defined as variability in the speaker's acceptability judgments on these constructions. The internalized grammatical knowledge of native speakers may be regarded as consisting of an indefinite number of acceptability hierarchies, ranging from a determinate core to an indeterminate periphery. The notion of acceptability hierarchy was found to be compatible with recent theoretical arguments put forward within generative grammar in favour of continuity between a UG-specified core and a UG-underspecified periphery.

Furthermore, it was claimed that the interlanguage grammars constructed by adult second language learners are affected by indeterminacy to a much greater extent than native grammars because of their inherent instability. Interlanguage indeterminacy does not necessarily decrease or disappear at higher levels of proficiency. Even the competence of near-native speakers may be characterized by incompleteness with regard to certain constructions, manifested in variable or inconsistent acceptability judgments. If near-native speakers come to attain acceptability hierarchies similar to those acquired by native speakers, then their intuitions on peripheral constructions will also be indeterminate (and their judgments will be variable), because of the intermediate grammaticality of these constructions, rather than because of incomplete knowledge representations. Indeterminacy due to incompleteness may be empirically indistinguishable from indeterminacy due to intermediate grammaticality.

These concepts represent the foundations of this thesis. Our aim is to investigate the development of linguistic intuitions in non-native grammars and to compare variation and indeterminacy in native and near-native acceptability judgments. The accomplishment of this aim requires the following elements:

(a) an area of grammar that presents indeterminacy in the intuitions of native speakers. This is defined in Chapter 3.

(b) a cross-sectional sample of learners at different proficiency levels, ranging from low-intermediate to near-native.[22] The lower proficiency groups have to be identified on the basis of some objective measure of progress from level to level, while the criterion for the

selection of informants at the near-native level is native-like performance in spontaneous verbal interaction. A description of the samples of learners who participated in the empirical studies can be found in Chapters 5 and 6.

(c) an empirical method for the elicitation of data that is sensitive to variation in acceptability and allows its measurement.

The next chapter is devoted to the establishment of such an empirical method.

CHAPTER 2
MEASUREMENTS OF LINGUISTIC ACCEPTABILITY

2.1 The elicitation of acceptability judgments

Given the complexity of the intuitional process, the elicitation of acceptability judgments is a much more problematic task than it is usually assumed. Minimally, the following conditions have to be met:

(a) in order to obtain valid data, the effect of the informant's internalized grammatical competence has to be disentangled from extra-grammatical factors that may interfere with it. This means that the researcher has to control for as many of these factors as possible.

(b) in order to obtain reliable data, the experiment should be replicated with the same subjects in the same experimental conditions.

(c) in order to obtain indicators of the status of acceptability hierarchies in the informant's competence, it is crucial to capture both determinacy and indeterminacy - and particularly degrees of indeterminacy - in linguistic intuitions.

(d) in order to provide an account of variability in judgments that can be legitimately generalized, it is necessary to find statistically valid ways of making comparisons on measurements of linguistic acceptability.

All the four conditions above are conspicuously overlooked in linguistic research. The linguistic intuition, while still the most often consulted behaviour in the linguistic literature, has not given rise to quantitative models comparable to those adopted in other experimental fields. The elicitation of acceptability judgments continues to take the form of an informal consultation of colleagues (often of the same theoretical persuasion), who are asked to say whether they like a sentence or not. And while it is now generally accepted, both inside and outside the framework of generative grammar, that judgments about the acceptability of sentences are relative (i.e..'more' or 'less') rather than absolute (i.e. 'yes' or 'no'), notations of varying acceptability like '**', '*', '??' and '?', which abound in the syntax literature, make very little advance on an all/none scale. The problem lies not so much in the limited number of symbols used (because more stars and questions marks could be added to the scale) but rather in the fact that such a scale is an *ordinal* measurement. It tells us that a sentence

classified '?' is more acceptable than a sentence classified '??', but gives no information about the difference in acceptability between the two sentences: whether this difference is of the same size, or smaller, or larger, than the difference between sentences classified '*' and '**'. For this reason, the scale cannot tell us whether linguistic predictions are associated with large or small differences in acceptability or, more crucially, with important or trivial factors in the informants' internalized linguistic competence.
Moreover, restricting the measurement of intuitions to an ordinal scale prevents our doing any interesting analysis of the results: biological and physical laws, inferential statistics, and even the most useful accounts of the variability of behaviour all demand at least an interval scale, one in which successive points are separated by equal intervals and the measurement can be subjected to arithmetic operations.

These problems affect both studies of native judgments and studies of non-native judgments, but given the inherent instability and indeterminacy of interlanguage grammars, they require more urgent attention in second language acquisition research. Is it possible to turn the elicitation of linguistic intuitions into a properly designed experiment?

The following sections will examine the characteristics of the most common elicitation techniques used in linguistic research and will discuss their shortcomings. Although the main focus will be on non-native judgments, most of the arguments that will be made apply to native speaker judgments as well. An alternative will then be suggested which borrows from the domain of psychophysics: magnitude estimation. It will be shown that the magnitude scaling of acceptability judgments is, in principle, feasible and allows us to approach variability of judgments directly.

2. 2 Reliability criteria

We saw earlier how the reliability question concerns the consistency of judgments, both within- and between-individual informants. There are two basic possibilities:

(a) inconsistency may be an artifact of the experiment (because of an inadequate test design, poor test items, etc). In this case, inconsistency may disappear if informants are given a properly designed test;

(b) inconsistency may be an effect of the inherent indeterminacy of the grammatical construction investigated. In this case, successive replications of the test would continue to produce inconsistent results.

Only the latter type of inconsistency is relevant here. However, we saw earlier that a structure may be indeterminate in the speaker's intuitions for a variety of reasons, which can be subsumed under the two basic notions of (a) developing knowledge, and (b) intermediate grammaticality. The former of course tends to characterize non-native judgments, whereas the latter may affect both native and non-native judgments. Replication of the test, as the reliability criterion most often recommended, is not always applicable to non-native intuitions, particularly if they present developmental indeterminacy. Greenbaum (1977), for example, proposes four reliability criteria for native speakers' intuitions, three of which are concerned with intra-subject consistency and one with inter-subject consistency. Let us examine them briefly:

(a) replication of the same test with different subjects belonging to the same speech community. This leaves us with the problem of defining 'the same speech community'. While in the case of native speakers this usually means 'other native speakers of the same language', is it more problematic to identify the speech community that non-native speakers would belong to. However, it is often assumed that learners at the same proficiency level, or from the same language background, share common characteristics and can therefore be regarded as members of the same group.

(b) replication of the same test with the same subjects after a lapse of time. What needs to be defined in this case is the optimal length of time between two successive administrations of the test. The interval cannot be too short if one wants to avoid producing a learning effect in the subjects. When working with second language learners, the interval cannot be too long since interlanguage grammars are in constant evolution. If learners are continuously exposed to the L2, even a short interval between two testing sessions may produce different results.

(c) replication of the same test with the same subjects, using different but equivalent materials. As Greenbaum points out, different lexical versions of the same sentence type should be built into experiments as a matter of course. Inconsistent responses to lexical versions may then indicate either unreliable subjects, or the effect of irrelevant syntactic or lexical differences among sentences. Equivalence among lexicalizations cannot be taken for granted (except perhaps for the most common constructions), since theoretically identical sentences may turn out to be different in the informants' perception (see an example in Bradac et al. 1980): appropriate statistical tests should be applied to control for the intrusion of irrelevant lexical differences. In the case of non-native informants, particularly at low proficiency levels, there is the additional problem

that lexicalizations may be judged differently because of ignorance, either of the lexicon, or of other irrelevant syntactic aspects of the sentences.

(d) replication of the test with the same subjects and the same test materials but using different kinds of measurements. The obvious question in this case is whether different measurements actually elicit the same aspects of informants' judgments or, in other words, whether they are equally valid.

We will now turn our attention to the last question in some detail. The discussion in the following sections will first focus on general principles and methods for the scaling of stimuli, particularly on (a) the type of scales for responses, and (b) the types of responses required of informants. It will then examine how these principles apply to the elicitation of acceptability judgments.

2.3 Types of measurement scales

There are four basic types of scales on which subjects are asked to respond in experiments concerned with the measurement of stimuli:

(a) category scales
(b) ordinal scales
(c) interval scales
(d) ratio scales

Category (or nominal) scales partition a given dimension into a set of mutually exclusive subclasses. The assignment of items to any one subclass involves the relations of equivalence and difference; all items assigned to the same subclass must be equivalent with respect to the property that defines that subclass and different with respect to items assigned to other subclasses. The latter relation can be represented by means of the symbol $\neq$, as in '$x \neq y$'). A category scale of acceptability would have a number of mutually exclusive subcategories defined, for example, as 'right' or 'wrong'. The labels that define the various subcategories may be interchanged, provided that this is done consistently : this operation does not affect the essential information in the scale.[1]

Ordinal scales imply that the items in one category are not only different from the items in other categories but also stand in some relation to them. Thus, if responses are expressed on an ordinal scale, items assigned to a category are *greater* (i.e. more preferred, more

difficult, higher, etc.) than sentences assigned to another category. This relation can be represented by means of the symbol >, as in 'x > y'.
An ordinal scale of acceptability would be one that requires informants to rank a set of sentences from most to least acceptable: such a scale would convey the information that sentence x is more (or less) acceptable than sentence y but it would not provide the interval between x and y. Nunnally (1967) mentions three other kinds of methods based on ordinal estimation: *paired comparisons*, in which subjects are required to rank stimuli two at the time in all possible pairs; *constant stimuli*, in which a standard stimulus is paired with each member of a constant set of stimuli; *successive categories*, in which subjects are required to sort a collection of stimuli into a number of distinct categories, which are ordered with respect to a given attribute.

Interval scales have all the properties of ordinal scales but in addition they specify the distance between any two points on the scale. An interval scale of acceptability, for example, would therefore provide not only the information that a given sentence is more acceptable than another, but also - and crucially - how much more acceptable it is. This relation can be represented as 'x - y = z'. Interval scales are characterized by a common and constant unit of measurement which attributes a number to all pairs of items on the scale. The unit of measurement and the zero point are arbitrary. One method of interval estimation mentioned by Nunnally (1967) consists of giving subjects two stimuli of unequal intensity and asking them to select a third stimulus which is halfway between the two in terms of a specified attribute.
The application of interval scales to non-metric continua often relies on the assumption that the informant's response of 'yes' to any one item is exactly equivalent to other affirmative responses to other items. Creating an interval scale for acceptability, for example, would imply that the response 'acceptable' referred to a sentence is exactly equivalent to the response 'acceptable' to any other sentence.

Finally, ratio scales have all the characteristics of interval scale but in addition have a true zero point at their origin. The ratio of any two points on such scales is independent of the unit of measurement. This type of relation can be represented as 'x/y = z'. On a ratio scale all four of these relations are possible to attain: (a) equivalence, (b) greater than, (c) known ratio between two intervals, and (d) known ratio of any of two values on the scale (Siegel and Castellan 1988: 31). Real ratio scales are difficult to find, or to create, outside the physical sciences.

2.3.1 Admissible operations with different measurement scales

The four types of measurement scales that we have just discussed differ in terms of the type of statistical manipulation of data that they allow. Measurement can be regarded as the process of assigning numbers to observations in order to represent quantities of attributes, so that it is possible to perform mathematical and/or statistical operations on them. The type relations that characterize the assignment of items to points on a scale define and limit the type of operations that can be applied on the data. The permissible operations must be the ones of the numerical structure to which the particular scale is isomorphic (Siegel and Castellan 1988: 32). This requirement separates nominal and ordinal scales, on the one hand, and interval and ratio scales on the other. Only interval and ratio scales are isomorphic to the structure of arithmetic: more precisely, it is the differences between values in an interval scale that are in such an isomorphic relation, whereas in a ratio scale it is the numerical values themselves. It follows that only interval and ratio scales permit the application of parametric statistical tests on the data, since these tests require the operations of arithmetic (i.e. means and standard deviations) on the original scores. Nominal and ordinal data do not satisfy the parametric assumptions because the intervals between successive classes are not equal: therefore the arithmetic operations implied by parametric statistics do not have substantial meaning when they are performed on such data.[2]

2.4 Types of responses: absolute vs comparative

The second important distinction that we need to consider is whether informants are required to produce an absolute judgment on each sentence separately, or to make relative judgments among different sentences. The first type of test involves a *rating* response, and the second type involves a *ranking* response. This distinction is clearly related to the choice of a particular type of measurement scale, as will be seen later.These two fundamental types of responses, and the corresponding scales of measurement, present different properties and, when applied to acceptability judgments, may yield qualitatively and quantitatively different patterns of results.

2.4.1 Absolute judgments

The most common measurement technique used in linguistic research is one or another form of *category scaling* in which informants rate a sentence or express an absolute judgment by choosing one of a fixed number of options. Typically, informants are presented with a binary choice: 'correct' vs 'incorrect', or 'good' vs 'bad', or 'acceptable' vs

'unacceptable', etc. Alternatively, judgments can be expressed by choosing among categories on a scale that has more than two points.

In general terms, this kind of measurement has a number of serious weaknesses. Category scales compress the informants' judgments into a limited number of options, so that sentences 'more or less' alike are placed into the same category. The resolution of the categories, that is, the proportion between the true range of stimuli and the fixed range points on the category scale, is usually poor, and at its worst in binary scales. Inevitably, the more limited the resolution of the scale, the more constrained the informants' responses are, and the more likely the researcher is to get uninformative data at best, and invalid responses at worst.[3]

But category scales present another major defect, arising specifically from the nature of the internalized competence that they are supposed to test. As we saw earlier, native speakers may not be able to express absolute judgments on the indeterminate areas of their native language. Given the pervasiveness of indeterminacy in interlanguage grammars, it is apparent that rating scales, especially in binary form, are even more inadequate for the investigation of non-native competence. Let us examine in some detail why judgments obtained by means of a dichotomous scale may provide inaccurate or deceptive information about the learner's state of competence.

Consider the case in which learners are asked to produce an absolute judgment ('correct' vs 'incorrect') on a construction that is not - or no longer - determinate in their interlanguage grammar. Any sentence exemplifying that construction with be marked as either correct or incorrect without having such a status in the interlanguage grammar: the learner's choice will be random. If there are different tokens of the same test sentence type, judgments may or may not be inconsistent. If they are not, the researcher is left with the deceptive impression that the construction is determinate in the informant's mental grammar. But if judgments are inconsistent, there cannot be any unambiguous interpretation of such inconsistency: no way of deciding whether it can be traced back to random choice, for example, because the learner has no representation of the structure in question, or whether it is due to intermediate or advanced indeterminacy (which, as we saw earlier, may be caused by the temporary coexistence of old and new rules in the learner's interlanguage grammar).[4] The validity of judgments obtained through binary rating scales is therefore highly dubious.

2.4.1.1 Three-point scales and the certainty dimension

The fundamental inadequacy of binary rating scales is not much improved if binary scales are replaced by three-point scales. The addition of a third category labelled 'not sure' or 'in between' to a dichotomous scale would seem at first sight to solve the problem of learners' random choice and consequent uninterpretability of the consistency factors. It might be argued, in fact, that because of the presence of such a category, informants would not be forced to produce an either/or judgment: they would be free to express their uncertainty by channelling their responses to the neutral category. In practice, however, the likelihood of obtaining invalid results is still quite high.

The first problem lies in the definition of the middle category, which can be chosen either because of a state of psychological uncertainty of the informant, or because of a state of intermediate grammaticality of the structure in question. In theory, it is possible for a sentence to be judged completely acceptable but with less than full confidence or, conversely, a sentence may be judged as being of intermediate grammaticality with complete confidence. Although in practice acceptability and certainty tend to be related, in the sense that informants choose the extremes of a scale with more confidence than when they choose intermediate points, this possibility cannot be ruled out, essentially because *certainty* and *acceptability* are two quite distinct and independent dimensions. This is suggested by a number of studies where the two factors were subjected to separate analyses. A study by Yule, Yana and Tsuda (1985) explored the relationship between correctness in answering and accuracy in judging the correctness of the answer chosen; the results reveal that subjects were more sure of their correct answers than of their wrong answers, but two other combinations emerged: non-confident correct answering, and very confident wrong answering.
Sorace (1985a), refining a procedure originally suggested by Kohn (1982), was able to calculate separate scores for consistency and certainty in interlanguage judgments, and concluded that the two values do *not* follow parallel developmental curves: learners are more confident at the initial stages of the acquisition process, regardless of the consistency of their responses, but their certainty decreases dramatically at intermediate stages, while the consistency of their judgments improves.

Decisions about how to label the middle category are therefore not trivial, since they may bias responses in either direction: if the category is labelled 'not sure' or 'don't know', informants may feel more inclined to interpret it as an admission of uncertainty or ignorance, whereas if it is named 'in between', they might use it only if they feel that, for some reason, a particular

sentence is neither acceptable not unacceptable. In practice, however, there is no guarantee that a label like 'in between' may not be interpreted in the same way as 'not sure': this suggests that intermediate categories on a rating scale are always difficult to interpret correctly, both from the point of view of the informant and from that of the experimenter. Furthermore, native and non-native speakers may have different preferences and biases, since non-native speakers are more likely to feel uncertain about, or to lack knowledge of, particular constructions; they are also less likely, at least up to an advanced proficiency level, to have the degree of linguistic sophistication necessary to consistently judge a construction as having intermediate grammaticality. Finally, personality factors may override genuine acceptability judgments: some learners may feel reluctant to choose the middle category (regardless of the label attached to it) because of a fear of revealing their uncertainty; other learners may be inclined to choose it most of the time in order not to commit themselves to a definite judgment.[5]

A possible solution may consist of including a *separate* scale for certainty, so that informants are asked to express both their acceptability judgments and the degree of confidence with which their judgments are produced. This would be conceptually justified, since, as we saw earlier, learners' perception of their own knowledge may be very different from their actual knowledge (see Klein's (1985) distinction between confirmation and criticalness of rules). The difficulty in this case lies in ensuring (through clear instructions and examples) that acceptability and certainty are perceived as distinct by the informants: giving certainty judgments involves providing a kind of 'meta-evaluation' about one's own judgments of acceptability, which may be a very demanding task to some informants.

Finally, scales including more than three points are statistically more reliable, and have a better resolution. Depending on how they are used, these scales may yield ordinal information[6] (see Nunnally 1967) but they do not solve the problem of how to interpret intermediate points. Subjects may find it difficult to maintain a stable and consistent criterion in their use of middle categories, unless these are clearly labelled and the labels are defined unambiguously. If they are, however, the experimenter imposes an additional constraint on the informants' responses. For category scales, the alternatives seem to be only two: either pre-determining the range and type of judgments, or leaving the informants free to interpret categories in their own way, with the obvious risk of obtaining substantially different interpretations.

To conclude, category acceptability scales are oblivious to the true range of stimuli and the subjective range of judgments (Lodge 1982). They are insensitive to indeterminacy or

uncertainty in judgments, or they conflate the two dimensions. They deny access to powerful statistical treatments (ANOVA, regression, etc) because they are based on a nominal (or at best ordinal) level of measurement, which involves an arbitrary number of categories and the arbitrary assignment of numbers to categories. In short, category scales may result in (a) misclassification of sentences and (b) loss of information .

2.4.2 Relative judgments

Ranking measurements are essentially comparative and require the expression of *relative* (not absolute) judgments. With ranking scales, sentences are typically presented in groups of two or more, and the informants responds to the 'more' or 'less' of acceptability.
It is a fact that people are better at making relative judgments than at expressing absolute judgments (see Nunnally 1967 on this point). This is probably because most judgments in life are inherently comparative; even when giving absolute judgments, subjects tend to relate their responses to (a) stimuli of the same kind that they have experienced in the past, and (b) the range of stimuli in the set presented.

As we saw earlier, the most straighforward way of obtaining relative judgments is to present informants with a set of sentences and ask then to rank them from 'most acceptable' to 'least acceptable' (or vice versa). The reliability of ranking tests involving more than three sentences is not clear, although experiments with native speakers have used sets of 6 (Snow and Meijer 1977) or even 11 sentences (Mohan 1977). Learners at low proficiency levels should not be expected to handle large sets of sentences with confidence.

An alternative method is that of *successive categories*, in which subjects are instructed to sort a collection of sentences - each individually printed on a separate card - into a number of distinct piles or categories, which are ordered with respect to their degree of acceptability. Unlike the procedure of ranking sets of sentences, this leaves informants free to manipulate the sentences and to move them around while distributing them into piles, without being tied to the original order of presentation. Variants of this method have been used in research on the structure of the mental lexicon (Miller 1969) and in research on native language transfer in second language acquisition (Kellerman 1978)[7]. This is one of the methods employed in the present study, as will be seen later.

All these methods allow for the possibility of tied ranks if two or more sentences have the same degree of acceptability. Depending on whether they are timed or not, these methods leave ample margin for correcting one's initial response.

Ranking scales present the following advantages over rating scales:

(a) The resolution of this type of scales is less constraining, since the only limit to the number of response categories that informants can use is given by the number of sentences: in principle, each sentence can be assigned to a separate category.

(b) They have greater psychometric plausibility , since they explicitly ask for relative judgments.

(c) They are more suitable to the purpose of capturing indeterminacy in acceptability, since informants are not forced to produce categorical judgments on sentences that are indeterminate in their competence.

However, most ranking methods (including the ones mentioned so far) generate ordinal measurements, and are therefore vulnerable to the same kind of statistical criticism as rating scales. Unless the experimenter instructs the subjects to mark the distance between successive categories, the result will be a simple rank-order. Although there are ways of obtaining interval responses from methods of successive categories (see Nunnally 1967; Lodge 1981 for a discussion), these imply assigning numbers to categories in a post-hoc fashion and in an essentially arbitrary way.
Is there an alternative to nominal or ordinal scales of acceptability? Can acceptability be measured on an interval scale?.

2.5 Magnitude scaling

> "[Refusing to use linguistic intuitions as data] would eliminate linguistics as a discipline, just as surely as a refusal to consider what a subject senses or perceives would destroy psychophysics....In both cases, [i.e. linguistics and psychophysics] we are trying (though in very different ways) to find a basis for intuitive judgments. In both cases, furthermore, the difficulty of obtaining reliable and relevant reports is quite apparent."
> (Chomsky 1961: 225)

The following sections will explore the parallel between linguistics and psychophysics in some detail. The questions that will be addressed is: can acceptability judgments be treated in the same way as judgments in a psychophysical experiment?

The primary aim of psychophysical investigations is to discover and describe the relationship between the objectively determined physical dimensions of stimuli and the subjective estimates of the magnitudes of those dimensions: psychophysical methods are essentially concerned with measuring the discrepancy between the informants' judgments and the

objective physical states of affairs about which judgments are made. As we saw earlier (see chapter 1, section 1.1.1.5), there is no such a thing as an objective measurement of linguistic acceptability in the mental representation of knowledge: there are linguistic theories, which make certain linguistic predictions about the possible outcome of an acceptability judgment test. It follows that the correctness of an acceptability judgment cannot be ascertained, because this would involve specifying the linguistic analogue to the light meters or meter sticks that are used in assessing the correctness of subjective judgments in a psychophysical experiment.

In this respect, linguistic acceptability is more similar to other socio-psychological dimensions, like prestige of occupation, seriousness of crime, etc, which lack objective metrics, than it is to physical dimensions like brightness, weight and length. Does this fundamental difference preclude the extension of the psychophysical paradigm and methodology to non-metric continua? The answer is negative: it has been shown by numerous studies (see Lodge 1981 for a review) that social opinions *can* be subjected to the same quantitative methods and analyses as psychophysical judgments/sensations. In particular, it has been demonstrated that the *magnitude scaling* approach can be successfully employed to validate social scales, thus providing a quantitative and powerful measurement of social opinions. This suggests that scales obtained though judgments on social variables obey the same laws as judgments obtained through judgments on sensory variables.

The extension of magnitude scaling to the domain of non-metric continua has not so far been concerned with linguistic acceptability. One of the purposes of this thesis is to investigate the validity of psychophysical methods in this domain. This requires, first of all, examining the way magnitude scaling methods are employed in psychophysics.

2.5.1 Magnitude Estimation

The simplest form of magnitude estimation requires the informant to associate a numerical judgment with a physical stimulus. It is therefore a form of cross-modality matching (Stevens 1971). Other versions of the method involve matching a different physical continuum to the one the stimuli belong to: so, for example, the relative brightness of different lights is estimated by the force of hand-grip on a handle (itself measurable by the experimenter in dyne/cm^2). The number continuum is often regarded as the most convenient to use because it is known to the informants and thus does not need to be introduced to them.

In numerical magnitude estimation, informants are presented with a series of stimuli of unequal magnitudes, one at a time in random order, and are asked to assign a number (the modulus) to the perceived magnitude of the first stimulus (the standard), and then successive numbers to the perceived magnitude of stimuli in proportion to the modulus. Prototypical instructions read:

> You will be presented with a series of stimuli in irregular order. Your task is to tell how bright/loud/long they seem by assigning numbers to them. Call the first stimulus any number that seems appropriate to you. Then assign successive numbers in such a way that they reflect your subjective impression. For example, if a stimulus seems 20 times as bright/loud/long, assign a number 20 times as large as the first. If it seems one-fifth as bright/loud/long, assign a number one-fifth as large, and so forth. Use fractions, whole numbers, or decimals, but make each assignment proportional to the brightness/loudness/length as you perceive it.
> (Stevens 1971: 428)

Experiments using numerical estimation have repeatedly shown that:

(a) informants are capable of using numbers to make proportional judgments of stimulation levels for virtually any sensory attribute.

(b) when the numerical estimates of the perceived strength of sensory stimuli are log-transformed, averaged, and plotted directly against the objectively measured physical values of the stimuli on log-log coordinates, the points typically fall along a straight line.[8] The law underlying this linear relationship between objective measurements and subjective perceptions can be summarized as follows: *equal stimulus ratios produce equal subjective ratios*, or, put in a different way, subjective judgments are approximately proportional to physical values.

This generalization is known as the Power Law, which may be represented as

$$\psi = R = kS^b$$

where ψ is the subjective magnitude, R is the magnitude of the response, S is the magnitude of the stimulus, b is the exponent which characterizes the relationship, and k is a constant of proportionality that depends on the unit used (Lodge 1981: 13; Stevens 1966: 530).

By reflecting ratios between judgments of stimuli, magnitude estimation provides more than interval measurement, because ratios depend not only on the intervals between measured points but also on the distance of each from the origin. Similar results have been obtained

where magnitude estimation was applied to linguistic perception of stimuli for which some objective interval scale is available: similarity of syllables from different languages (Takefuta, Guberina, Pizzamiglio, and Black 1986), speech rate (Green 1987), vowel roughness (Toner and Emanuel 1989), quality of synthesized speech (Pavlovic, Rossi, and Espesser 1990), speech intelligibility (Fucci, Ellis, and Petrosino 1990).

2.5.1.1 Experimental decisions in magnitude estimation

The experimenter has to make a number of crucial decisions in setting up a psychophysical experiment. Some of these decisions are controversial and have raised criticism because they may prejudice the validity of magnitude estimation. The following section will present the 'canonical' recommendations (grouped under major headings) on how to run a magnitude estimation experiment.

(a) *The standard.* Stevens (1971: 428) argues that "it is usually better not to designate a standard. The subject then remains free to choose his own modulus". This is one of the decisions that the experimenter has to make, and one that may affect the way informants match numbers to sentences. Notice that the instructions above do not *explicitly* encourage informants to use the first estimation as a reference. If the stardard is provided and kept visible for the whole length of the experiment, then obviously informants will use it as the reference; if the standard is not provided, then the first stimulus should, in principle, become the reference.
However, depending on the length of the sequence of stimuli presented, informants will tend to forget the first stimulus and the number assigned to it: the standard will therefore shift during the course of the experiment and the immediately preceding stimulus will become the reference (Poulton 1986: 183 refers to this effect as the "sequential contraction bias"). The way instructions are phrased may determine the course of actions that informants take when faced with a magnitude estimation task. Consider the following set of instructions, which are superficially very similar to the previous set:

> You will be presented with a series of ----- stimuli in irregular order. Your task is to tell how intense they seem by assigning numbers to them. Let the first stimulus be your reference. Give it any number that seems appropriate to you, keeping in mind that some of the stimuli will be more ---- than the reference, and others will be less ---- than the reference. Assign a number to each of the stimuli such that it reflects how much weaker or stronger it is compared with the first stimulus: the more ---- it is compared to the reference, the bigger your number response; the less ---- it is

compared to the reference, the smaller your number response. There is no limit to the range of numbers you may use.
(Lodge 1981: 7, adapted from Stevens 1975: 30)

These instructions place much more emphasis on the first stimulus being the reference. It is possible that informants who receive them will make a stronger effort to remember the first stimulus and compare each successive one to it.
A different way in which the instructions may actually bias responses is by the very fact of providing numerical examples: subjects have the tendency to use the numbers suggested in the examples. This and other effects have been identified as 'transfer effects' (Poulton 1986, 1989).

(b) *The first stimulus*. Another important decision to be made concerns the place of the first stimulus within the overall range of stimuli presented. If the first stimulus is closer to the lower extreme of the range, there is likely to be a floor effect in the informants' responses; if the first stimulus is closer to the higher extreme of the range, there is likely to be a ceiling effect. A solution is to choose the first stimulus from those in the middle region of the range, and this is in fact what Stevens recommends. This, however, contradicts the principle of random order of presentation; using different random orders in the presentation of stimuli should counterbalance the potential effect of the first stimulus.

(c) *Training informants*. Stevens (1971, etc) claims that there is no need to train the informants, because "since there is no right or wrong to the subjects' responses, it is not clear what would be meant by training." (1971: 429). He concedes, however, that the nature of the task may be clarified by having the informants matching numbers to an 'easy' continuum like line length before the proper experiment. Other researchers agree that some form of training is desirable, because the the act of matching numbers to the perceived strength of stimuli is unfamiliar to most people (Lodge 1981) or, perhaps more crucially, because the concept of proportionality is alien to many .

(d) *Number of judgments required*. No more than one or two judgments per stimulus should be required by each informant (Stevens 1971: 428). This is in order to prevent learning effect, and also to minimize certain biases due to range and spacing of stimuli (Poulton 1968, 1989).

(e) *Averaging means*. Averaging with magnitude estimation data is usually done by taking either the median or the geometric mean of the estimates. The geometric mean is the

arithmetic mean computed on variables that have been transformed into logs and then exponentiated back to the linear scale. This is necessary because different informants may use different units of modulus, which can determine an atypical, skewed distribution of scores. The logarithmic transformation tends to undo the skewness, i.e. to normalize the scores.[9]

2.5.1.2 Biases in magnitude estimation

Magnitude estimation has been the target of a number of criticisms (see Poulton 1989 for a comprehensive review). Most of these are concerned with the validity of the Power Law, that is, of the linear relation between subjective judgments and the stimulus magnitudes; therefore they do not necessarily extend to the application of magnitude estimation outside the domain of psychophysics. The argument underlying these criticisms is that experiments using magnitude estimations are subject to certain artifacts and biases, only some of which can be controlled for. The main source of bias is the fact that most informants tested in psychophysical experiments are untrained observers who are required to use unfamiliar units of measurement. However, most of the biases identified in the literature apply to versions of the magnitude estimation method in which the range, or the direction of responses available to the informants to make estimates is established by the experimenter: this happens, for example, when informants are provided with a modulus and instructed to give only ascending, or descending judgments. Most biases can be avoided by leaving informants free to choose their response range.[10]

2.5.1.3 Extensions of magnitude estimation to non-metric dimensions

We saw earlier that the main difference between physical and non-physical continua is that the latter lack objective metric properties. Estimates of loudness can be plotted in a graph on the y axis against objectively measured decibels of sound pressure on the x axis, and the relationship between the two can be appraised. But in the case of social-psychological stimuli, what is to be placed on the X axis?

The solution suggested by Lodge (1981) relies on the *cross-modality matching paradigm* developed by J.C. Stevens et al. (1960). To quote Lodge:

> "If the power law is valid and if the exponents derived from magnitude estimation are truly characteristic, then any *two* quantitative response measures with established exponents could be used to judge a sensory continuum and the validity of the derived ratio scale confirmed by

obtaining a close match between the theoretical and empirically obtained ratios *between the two response measures*."
(Lodge 1981: 28)

The extension of magnitude estimation to social-psychological variables is governed by the principle that "the empirically obtained ratio between response modalities when matched to social stimuli should *approximate* the ratio established for the same two response modalities when matched to physical stimuli." (Lodge 1981: 31). This involves the comparison between the exponents obtained from the estimates of socio-psychological stimuli through two response modalities and the exponents obtained from the same informants when they match the same two response modalities to metric stimuli. The application of magnitude estimation to non-metric continua thus implies three stages:

(a) in the first stage (*calibration task*), informants are given instruction and practice in making proportional judgments of metric stimuli by using two quantitative response modalities (i.e. line length estimation and numerical estimation), each used to estimate the stimuli properly belonging to the other. This stage gives the informants the opportunity to receive some training in using the concept of proportionality. It also provides the experimenter with a way of determining whether such a concept has been understood and accurately applied. When the geometric means of the two sets of judgments are plotted against each other, the result should be a straight line;

(b) in the second stage, informants use the same two response modalities to judge the intensity of the socio-psychological stimuli. The assumption here is that informants transfer the skills developed for the metric continua to the socio-psychological continuum. This assumption, of course, cannot be verified directly . If subjects are able to estimate the magnitude of non-metric stimuli as they could with interval continua, then the plot of the two magnitude estimates of the non-metric stimuli would again cluster around a straight regression line;

(c) in the third stage, the slope of the regression line obtained in the calibration session and that of the regression line obtained in the estimation of social stimuli are compared: in order to validate the magnitude scaling of social stimuli, the two slopes should not be statistically different from one another (see Lodge 1981 for a complete account of the experimental and statistical techniques required by the validation procedure).

2.5.2 Magnitude scaling of linguistic acceptability

We can now return to the question of the elicitation of acceptability judgments. Can magnitude estimation be used for this purpose?

If linguistic acceptability is a continuous non-metric dimension, and if acceptability judgments are subjected to the same laws as any other human judgment, then the answer is definitely positive. Magnitude estimation is a method that requires relative, or comparative judgments. Unlike other ranking procedures, however, it offers the following advantages:

(a) it produces at least interval scales, which can be legitimately subjected to parametric statistics for the analysis of variance.

(b) it places no constraints on the number and range of responses available to informants, who are thus able to express precise judgments without compressing them into a pre-determined scale.

(c) it thus allows us not only to capture variability and indeterminacy in judgments, but also to discover the differences among degrees of acceptability of sentences falling on an acceptability hierarchy.

In its canonical form, moreover, magnitude estimation is a timed procedure which encourages informants to respond before they can retrieve metalinguistic rules, and prevents them from changing their minds. It is therefore plausible to assume that magnitude estimation, unlike other ranking procedures, may tap 'tacit', rather than metalinguistic, knowledge.

To this list of advantages one can add convenience of use (particularly numerical magnitude estimation), and easy applicability to a large number of sentences.

We saw at the end of Chapter 1 that one of the elements essential to the investigation of indeterminacy in native and non-native linguistic intuitions is a method that can measure systematic variation in acceptability judgments. Magnitude estimation has all the requisites that we were looking for. There is no independent evidence, however, that this method will work, because there have been no previous applications of magnitude estimation procedures to this purpose.

This thesis represents the first large-scale application of magnitude estimation as a tool for the study of variation in native and non-native linguistic acceptability but, given its wider scope, it will consider only some of the questions raised in this chapter.

In particular, the experimental criteria discussed in 2.5.1.1 will be met as follows:

(a) the standard will not be designated, in order to leave informants free to choose their own moduli;

(b) the sentences will be presented in different random orders, so that potential biases induced by the position of the first stimulus within the range are counterbalanced;

(c) informants will receive explicit instructions and examples about how to provide ratio judgments of linguistic acceptability;

(d) geometric means will be used as measures of central tendency of the scores.

The questions of cross-modality matching and calibration of response modalities, however, will not be addressed. Our less ambitious, but more realistic aim will be to compare magnitude estimation with an untimed ranking procedure which is based on an ordinal level of measurement: card-sorting. Our weaker hypothesis is that if magnitude estimation is a viable method for the elicitation of acceptability judgments, then it should be as least as informative as card-sorting. Our stronger hypothesis is that, given that it is a more powerful method than card-sorting, magnitude estimation should reveal a more discriminating - and meaningful - pattern of acceptability. A confirmation of the stronger hypothesis will justify a more in-depth validation of magnitude scaling of linguistic acceptability.

A description of the experimental design underlying the application of magnitude estimation and card-sorting will be given in Chapter 6.

We will now focus on the area of grammar that will allow us to investigate variation and indeterminacy in linguistic acceptability: auxiliary selection in Italian.

CHAPTER 3
AUXILIARY SELECTION IN ITALIAN: AN OVERVIEW

3.0 Introduction

This chapter has two aims: first, to review some of the basic research concerning auxiliary choice, and to provide a unifying perspective on the different approaches in this area; second, to show that a comprehensive account of the phenomena related to auxiliary selection needs to focus on the interaction between the syntactic and the lexical-semantic dimensions. It will be finally argued that it is particularly the lexical-semantic dimension that allows an understanding of the evolution and acquisition of Italian auxiliaries.

The main theoretical accounts proposed in the literature generally differ with respect to whether they focus on syntactic configurational properties or on lexical-semantic characteristics of verbs. These are often presented as mutually exclusive explanations. Syntactic theories derive grammatical phenomena from a restricted number of abstract principles which are also responsible for other - seemingly - unrelated properties; they tend to explore the way particular verbs behave with respect to these properties. Semantic theories, on the other hand, tend to concentrate less on the internal consistency of the overall system than on the details of the individual phenomena; they focus on how classes of verbs that select different auxiliaries are defined (see Grimshaw 1987). This dichotomy in the research on auxiliary selection does not make the task of comparing theories a straighforward one. However, a careful appraisal of the different theoretical positions makes clear that syntactic and semantic approaches are much more compatible than they appear at first sight. Indeed, the underlying theme to this chapter is that one has to look at the interface between syntax and semantics in order to account for the complexity of the grammatical phenomena related to this area. More importantly, we want to show that it is ONLY by looking at lexical-semantic distinctions internal to verb classes that one can interpret the patterns of evolution of auxiliary use, both in historical terms and in terms of language acquisition.

The theoretical overview presented in this chapter will identify an area of systematic variation in the grammar of Italian. This will provide the basis for an empirical study , which will test the hypothesis that verb classes characterized in syntactic terms as internally homogeneous may not be perceived as such in the linguistic intuitions of

native Italian speakers. The working assumption will be that linguistic intuitions (both native and non-native) are sensitive to structural and semantic aspects of sentences, although not necessarily the same ones in the native and non-native cases. One of the issues addressed by this study concerns precisely the extent to which non-native speakers who begin learning a foreign language as adults are still sensitive to purely syntactic configurations (under the assumption that the latter have at least some degree of 'psychological reality' in native knowledge representations), as opposed to semantic aspects of grammatical phenomena. The prediction that will be subjected to empirical testing is that syntactic phenomena that lend themselves to semantic characterizations are more learnable than phenomena that do not.

This chapter is structured as follows:

(a) First, the research on syntactic aspects of auxiliary selection will be discussed with reference to the two main versions of the Unaccusative Hypothesis, namely the Relational Grammar version (Perlmutter 1978, 1989) and the Government-Binding (GB) version (Burzio 1986). The discussion will particularly centre on the GB account, which reveals subtle connections between unaccusativity and auxiliary choice not only in the unmarked context of the present perfect, but also in 'marked' cases, such as the 'restructuring' constructions originally studied by Rizzi (1982). The main focus of this first section will be Italian.

(b) Second, the syntactic analysis will be extended to French, and to a lesser extent to English. Some of the limitations of an purely syntactic approach will be pointed out.

(c) Third, research on lexical-semantic aspects of auxiliary selection will be presented. This section will cover:

   i. a concise summary of the historical evolution of auxiliaries from Latin to Romance;

   ii. semantic analyses centred on thematic aspects: Vincent (1982); Tuttle (1986); Keenan (1987);

   iii. semantic analyses centred on aspectual features: Parisi (1976), van Valin (1989, 1990), Centineo (1986)

   iv. research on the organization of argument structure and on the mapping from semantic representations onto syntactic representations, with reference to Rappaport and Levin (1986), Zubizarreta (1987), and particularly to Grimshaw (1991).

(d) Fourth, and centrally for the empirical part of this study, a proposal will be put forward for the analysis of unergative and unaccusative verbs in terms of acceptability hierarchies. It will be argued that verbs commonly subsumed under the 'unaccusative' and 'unergative' labels actually differ in status: distinctions can be made at the level of lexical-conceptual representation that are consistent with the overall pattern of diachronic and syncronic variation and allow specific predictions for language acquisition.

3.1. The Unaccusative Hypothesis and auxiliary selection

The fundamental insight (originally due to work by Perlmutter (1978)) that has opened up a new perspective on the behaviour of auxiliaries is that the class of verbs traditionally referred to as 'intransitive' is not homogeneous, but consists of two main sub-types: unaccusative and unergative. Unaccusative verbs are the ones that, although of the form NP-verb on the surface, are assumed to have an underlying structure where the NP is the direct object of the verb, as in (1) below:

(1) [e arriva Marco] --> [Marco$_i$ arriva e$_i$]

Unaccusative verbs contrast with unergative verbs which, although also of the form NP-verb on the surface, are assumed to have an underlying structure where the NP is the subject of the verb.

This difference forms the basis of the Unaccusative Hypothesis, formulated within the theoretical frameworks of Relational Grammar (Perlmutter, 1978; Perlmutter and Postal, 1984; Perlmutter, 1989) and Government- Binding (Burzio, 1986),details of which will be discussed in the following sections. Such differences can be provisionally schematized as follows:

(2)	UNACCUSATIVE	UNERGATIVE	
	i. Initial: V2	Initial: V1	(Relational Grammar)
	ii. $[_S[_{NP}$ e$]$ $[_{VP}$ V NP$]]$	$[_S$ NP$[_{VP}$ V$]]$	(Government-Binding)

The reason why a class of unaccusative verbs has been proposed is that the NP associated with such verbs shares a number of syntactic distributional properties with the direct object of transitive verbs which it does not share with the subject of

unergative verbs. Some of these properties are illustrated by the examples below from Italian and French:

(a) the possibility of partitive NE/EN ('NE/EN cliticization'), that is, the pronominalization of a quantified post-verbal subject in the form of the clitic pronoun NE/EN. This type of cliticization is systematically possible with direct objects of transitive verbs, and systematically impossible with unergative verbs:

(3) a. Ne ho visti molti (di bambini) — transitive
J'en ai vu beaucoup (d'enfants)
(I) of-them I have seen many (of children)
'I saw a lot (of children)'

b. Ne sono partiti molti (di turisti) — unaccusative
Il en est parti beaucoup (de touristes)
Of-them have left many (of tourists)
'Many (tourists) left'

c. *Ne hanno parlato due (di professori) — unergative
*Il en a parlé deux (de professeurs)
'Of-them have spoken two' (of professors)
'Two (professors) spoke'

(b) the use of adjectival participles, always possible with transitive verbs and impossible with unergative verbs:

(4) a. Il programma trasmesso ieri sera era interessante — transitive
L'émission diffusée hier soir était intéressante
'The programme broadcast yesterday was interesting'

b. Lo studente arrivato ieri è americano — unaccusative
L'étudiant arrivé hier est américain
'The student arrived yesterday is American'

c. *L'amico parlato stamattina... unergative
*L'ami parlé ce matin...
'The friend spoken this morning...'

(c) the use of participial absolutes, again allowed by transitive verbs and disallowed by unergatives:

(5) a. Arrestato dalla polizia, Paolo ha subito un lungo interrogatorio transitive
Arrêté par la police, Paul a subi une longue interrogation
'Arrested by the police, Paul underwent a long interrogation'

b. Caduto dal seggiolone, il bambino si è messo a urlare unaccusative
Tombé de sa chaise, le bébé s'est mis à hurler
'Fallen from his chair, the baby started to scream'

c. *Lavorato per trent'anni, Gianni andò in pensione unergative
*Travaillé pendant trente ans, Jean prit sa retraite
'Worked for thirty years, John retired'

The class of unaccusative verbs includes:

(a) verbs in regular alternation with a transitive counterpart , so that the subject of the unaccusative member of the pair corresponds to the direct object of the transitive member. This subset of verbs comprises both verbs in active form (e.g. *aumentare/augmenter* 'increase', *affondare/couler* 'sink'), and verbs in reflexive form (e.g. *bagnarsi/se mouiller*, 'get wet', *coprirsi/se couvrir*, 'cover up', *trasformarsi/se transformer*, 'turn into');

(b) inherently reflexive verbs (e.g. *fidarsi/se fier* 'trust', *pentirsi/se repentir*, 'regret', *suicidarsi/se suicider* 'commit suicide')

(c) verbs with no lexicalized transitive counterparts. These comprise the two subgroups of verbs without an unergative counterpart (e.g. *arrivare/arriver* 'arrive', *diventare/devenir* 'become'), and verbs with an unergative counterpart (e.g. *correre/courir* 'run', *volare/voler*, 'fly').

Unaccusative verbs with either transitive or unergative counterparts will be henceforth referred to as *paired* unaccusatives (see Perlmutter 1989), whereas verbs with no alternants will be referred to as *unpaired* unaccusatives.

Auxiliary selection in the present perfect tense has often been identified as one of the principal diagnostics of unaccusativity or unergativity: unaccusative verbs tend to select ESSERE/ETRE, whereas unergative verbs select AVERE/AVOIR. It is with respect to auxiliary selection that Italian and French most obviously diverge: while ESSERE is a necessary and sufficient condition of unaccusativity in Italian, ETRE is a sufficient but not a necessary condition of unaccusativity in French. A comparison between the two auxiliary systems will reveal that Italian unaccusative verbs as a class behave consistently , whereas French unaccusative verbs do not.

The next section will present the facts of unaccusativity and auxiliary selection in Italian. It will be followed by a summary of the Unaccusative Hypothesis in relational Grammar and in Government and Binding (which was originally formulated to account for Italian intransitive verbs). The French auxiliary system will then be introduced in section 3.5, where it will be shown how the Unaccusative Hypothesis accounts for some, but not all, of the phenomena related to auxiliary selection in French.

3.2 The Italian auxiliary system

Auxiliary selection is largely regular and predictable in Italian. While all transitive and unergative verbs take AVERE, ESSERE is systematically found with all unaccusative verbs, as well as with all passives and reflexives.

(6)	a.	Maria ha mangiato una mela	transitive
		'Maria has eaten an apple'	
	b.	Mia sorella ha viaggiato in treno	unergative
		'My sister has travelled by train'	
	c.	La legge è sostenuta dal governo	passive
		'The law is supported by the Government'	
	d.	Paolo si è rasato	reflexive
		'Paolo has shaved'	
	e.	I prezzi sono aumentati	paired unaccusative
		'Prices have increased'	(active form)
	f.	Il vaso si è rotto	paired unaccusative
		'The vase has broken'	(reflexive form)

g. Paolo è andato a casa — unpaired unaccusative
'Paolo has gone home'

h. Paolo non si è accorto di niente — unpaired unaccusative
'Paolo hasn't noticed anything' — inherently reflexive

All Italian unaccusative verbs, whether paired or unpaired, in active or reflexive form, satisfy all the tests of unaccusativity in (3), (4) and (5) above.

The predictability of auxiliary selection in Italian is further confirmed by the regular behaviour of paired unaccusatives with an unergative alternant: the unaccusative member of the pair takes ESSERE and satisfies the usual tests of unaccusativity, whereas the unergative member takes AVERE and does not satisfy these tests.

(7) a. Luigi è corso alla stazione — unaccusative
'Luigi has run to the station'

b. Alla stazione, di turisti, ne sono corsi molti
'To the station, of tourists, of-them have run many'

c. Luigi ha corso velocemente — unergative
'Luigi has run fast'

d. *Di atleti, ne hanno corso velocemente molti
'Of athletes, of-them have run fast many'

A small group of verbs, including the modals *dovere, potere, volere* and the aspectuals *cominciare* and *stare per*, optionally change their auxiliary AVERE to ESSERE when they are followed by an embedded clause containing an ESSERE verb:

(8) Maria ha/è dovuta partire
Maria has/is had to leave
'Maria had to leave'

As will be seen below in detail (see section 3.4.1.3), auxiliary change is related to other syntactic properties, such as the movement of a clitic originating in the embedded verb to the matrix verb. This 'clitic climbing' is also optional but requires obligatory change of auxiliary in the matrix verb from AVERE to ESSERE,

if the matrix verb independently requires AVERE . When there is no clitic movement, i.e. when the pronoun complement remains attached to the embedded verb, auxiliary change is again optional, and either ESSERE or AVERE is allowed .

(9) Mario ci è/*ha dovuto andare (a casa)
Mario there is/has had to go
'Mario had to go there'

Other verbs allowing clitic climbing are some unaccusative verbs of motion, such as *andare* and *venire*. Auxiliary change in sentences not involving clitic movement does not apply to these verbs, since they independently take ESSERE.

This description of the facts of auxiliary selection in Italian reveals a largely predictable system in which the distinction unaccusativity/unergativity and the choice of auxiliaries are closely related. Let us now turn to explaining what motivates such a close relation.

3.3 The Unaccusative Hypothesis in Relational Grammar

The early formulations of the Unaccusative hypothesis (Perlmutter 1978; Perlmutter and Postal 1984) differ rather substantially from more recent theoretical statements (Perlmutter 1989): the former posits an explicit link between grammatical relations defined in syntactic terms and (potentially universal) semantic roles, which goes under the name of 'Universal Alignment Hypothesis', whereas the latter abandons any attempt to characterize unaccusativity and its related properties in semantic terms, and therefore any universally, or at least cross-linguistically, valid definition of unaccusativity, resorting entirely to language internal syntactic evidence .
This section will provide a concise account of the syntactic bases of unaccusativity within a Relational Grammar framework,[1] which have remained constant. A detailed analysis of the semantic implications will be presented in section 3.8.

In Relational Grammar, the notions 'subject' and 'object' are theoretical primitives. Clause structure is represented in terms of a network of grammatical relations, or a relational network (RN). Crucial to the definition of 'syntactic level' is the concept of 'stratum', which is a set of all arcs with the same tail sharing some coordinate.

Nominals can bear different grammatical relations to a clause in different strata, so that one can say that a given nominal is a subject, or a direct object, only with respect to a particular stratum. The crucial feature of this theoretical framework is the assumption of a double level of syntactic representation, and particularly the existence of a level of initial representation (i.e. an initial stratum) distinct both from the final level (i.e. a final stratum) and from any thematic level (Perlmutter 1989: 99).

The traditional notions of transitivity and intransitivity are represented as follows:

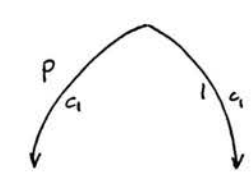
(10) a. A stratum is *transitive* if and only if it contains a 2-arc
b. A stratum is *intransitive* if and only if it contains a 1-arc and no 2-arc

The assumption underlying (10) is that there are strata containing a 1-arc and no 2-arc, but there are no strata containing a 2-arc and no 1-arc. The Unaccusative Hypothesis represents a challenge to the above assumption by distinguishing between two types of intransitive strata:

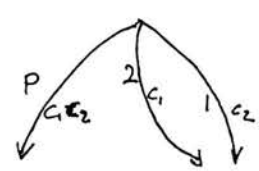
(11) a. A stratum is *unergative* if and only if it contains a 1-arc and no 2-arc
b. A stratum is *unaccusative* if and only if it contains a 2-arc and no 1-arc

The basic claim of the Unaccusative Hypothesis is that intransitive clauses divide into two groups depending on whether the initial stratum is unergative or unaccusative. The Unaccusative Hypothesis interacts with the Final 1 Law of Perlmutter and Postal (1983), which requires that the final stratum of basic clauses contains a 1-arc and therefore prevents final strata from being unaccusative. One way in which this can be prevented is by having the initial 2 of an initially unaccusative clause advance to 1: this is known as 'Unaccusative Advancement'.
Simplified representations of initially unergative and initially unaccusative clauses are illustrated in (12):

(12) a. Unicorns sleep

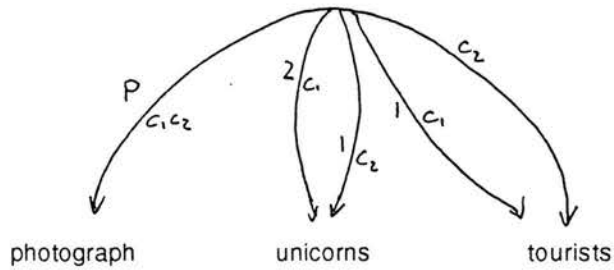


b. Unicorns exist



Unaccusative Advancement also applies to produce passive clauses: in both cases, a nominal advances from 2 to 1, but the difference is that in passive clauses the stratum in which the 2 advancing to 1 heads a 2-arc also contains a 1-arc:

(13) a. Unicorns were photographed by tourists



Within this framework, all the syntactic phenomena associated with initial unaccusativity (i.e. the possibility of participial absolutes, of participial adjectives, etc.) are shown to be sensitive to 2hood. For instance, the general condition for NE-cliticization is the following:

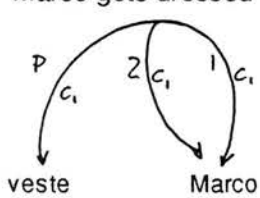
(14) only a nominal heading a 2-arc can be the source of partitive NE.

The generalization captured by this condition unites unaccusative clauses, whose initial stratum has a nominal heading a 2-arc, and transitive clauses, which have a 2-arc.

Essential to an explanation of auxiliary selection is the theoretical construct known as 'Multiattachment", according to which a given nominal can head more than one arc with the same tail in the same stratum. The meaning of this construct overlaps to a considerable extent with the semantic notion of 'coreferentiality' and therefore, in Perlmutter's view, it makes such notion unnecessary. This can be seen in the representation of the following reflexive clause in which the nominal 'Marco' heads both a 1-arc and a 2-arc:

(15) Marco si veste

'Marco gets dressed'



Multiattachment applies only to initial strata, because a nominal cannot head more than one arc in the final stratum of any clause. Multiattachment has to be 'resolved' so that no nominal heads more than one arc with the same tail in the final stratum. (Details of the resolution of multiattachment are irrelevant in the present context and will be omitted).

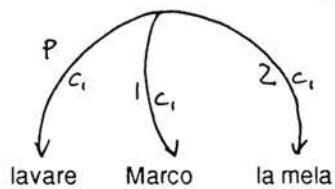
The rule for auxiliary selection relies on the notion of multiattachment as follows:

(16) If a clause contains a nominal heading both a 1-arc and either a 2-arc or a 3-arc , then the clause b requires ESSERE. Otherwise it requires AVERE.

This statement covers all the basic cases of auxiliary selection, including both transitive and intransitive clauses, as can be seen in the examples in (17) below, where (a) is transitive, (b) is reflexive, (c) is unergative, and (d) is unaccusative:

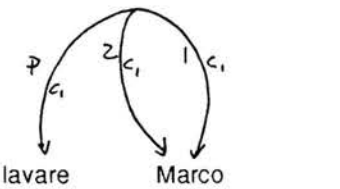
(17) a. Marco ha lavato la mela

'Marco has washed the apple'



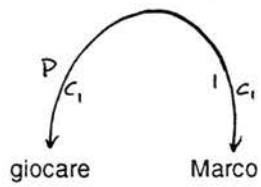
b. Marco si è lavato

'Marco has washed himself'



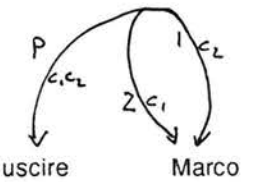
c. Marco ha giocato

'Marco has played'



d. Marco è uscito

'Marco has gone out'



Notice, however, that although this apparatus makes the correct predictions about the distrubution of auxiliaries, it does not *explain* why Multiattachment requires ESSERE and not AVERE. As we shall see later in the chapter, the Government and Binding version of the Unaccusative Hypothesis is vulnerable to the same criticism.

3.3.1 Summary

Recent formulations of the Unaccusative Hypothesis in Relational Grammar emphasize the independence of the syntactic analysis from any semantic or thematic representation, and its ability to account for auxiliary selection and other grammatical properties related to unaccusativity without any resort to semantic notions. This, together with the assumption of a double level of syntactic representation, makes the Relational Grammar account similar to the Government-Binding one, as will be seen in the next section. However, there is a major difference. In terms of Relational Grammar, the autonomy of the syntactic analysis implies abandoning the idea that the Unaccusative Hypothesis has cross-linguistic validity. In terms of Government and Binding, the syntactic analysis that accounts for Italian can be naturally applied to other languages, where surface differences are evidence of parametric variation with respect to invariant principles. The two frameworks ultimately differ on the question of the 'psychological reality', and therefore of the universality, of syntactic constructs.

3.4 The Unaccusative Hypothesis in Government and Binding[2]

Burzio's encyclopedic book *Italian Syntax: a Goverment Binding Approach* provides a principled explanation of the systematic behaviour of Italian auxiliaries. The central idea in Burzio's account of auxiliary assignment is borrowed from Perlmutter and recast in Government-Binding terms, namely the identification of a class of ergative verbs in addition to the traditional classes of transitive and intransitive verbs. Given the perfect equivalence of the meaning of two terms 'unaccusative' and 'ergative', and of 'unergative' and 'intransitive' the present study will continue to employ Perlmutter's original terms 'unaccusative' and 'unergative' (which have the advantage of leaving 'intransitive' available to define the whole set of unaccusative and unergative verbs in opposition to transitive verbs).

(18) a. Transitive — Luigi compra il giornale
'Luigi buys the newspaper'

b. Unergative — Luigi parla
'Luigi speaks'

c. Unaccusative — I prezzi aumentano
'Prices increase'

Cases like (18c) are derived via NP-movement from a corresponding transitive structure:

(19) a. [e] aumentare i prezzi
increase prices
b. [I prezzi$_i$] aumentano t_i
Prices increase

Although motivated by a different set of theoretical principles, this derivation is perfectly analogous to the notion of 'Unaccusative Advancement' in Relational Grammar.

Burzio refers to pairs like (19) as "AVB/BV"surface structure pairs, where V is a verb and A,B are noun phrases. In these pairs, the inverted subject (henceforth i-subject) is semantically related to a direct object, in much the same way as in passive sentences. Compare the two alternations below:

(20) i. a. Gli studenti hanno letto molti libri
'The students have read many books'

b. Sono stati letti molti libri
Are been read many books
'Many books have been read'

ii. a. I nemici affondarono due navi
'The enemy sank two ships'

b. Affondarono due navi
Sank two ships
'Two ships sank'

In each of the above pairs, the phrase underlined is the patient in each (b) case, just as it is the patient in the (a) case. Moreover, in distributional terms, it is the same class of NPs that can occur in both members of each pairs. In both cases, S-structure subjects are D-structure direct objects.

The two verbs *aumentare* in (19) above have identical subcategorization frames and differ by exactly one lexical parameter, namely whether or not they assign a theta-role to the subject position: the verb in (19b) does not assign such theta-role, and therefore no argument will be found, whereas the verb in (19a) does, so that the D-structure representation requires the presence of an argument. In fact there is no need to assume two different verbs, if one assumes only one verb that assigns subject theta-role optionally. It should be noticed that the possibility of non assignment of theta-role to the subject position is established independently for various verbs taking sentential complements, such as Raising verbs:

(21) a. [e] seems [John to leave] --> John seems to leave
b. [e] seems [that John left] --> It seems that John left

The possibility of assigning subject theta-role can therefore be seen as a parameter that varies among verbs which are subcategorized for NP objects, just as it does for verbs subcategorized for sentential complements. Since the variation is determined by lexical -not syntactic - factors, one would not expect pairs like (19) above to appear with full productivity, and this is what happens in Italian. On the one hand, there are verbs like (22a), lacking the counterpart (22b):

(22) a. Maria scrive la lettera
'Maria writes the letter'

b. *La lettera scrive
'The letter writes'

On the other hand, there are verbs like (23a), which appear in a D-structure frame "[e] V NP" and lack a lexicalized transitive counterpart.

(23) a. [e] arrivare molti studenti --> Molti studenti arrivano
arrive many students 'Many students arrive'

b. *Francesco arriva molti studenti
Francesco arrives many students

Verbs like (22) would be characterized as AVB/*BV, whereas verbs like (23) - which correspond to Perlmutter's unpaired unaccusatives - would be cases of *AVB/BV.[3]

The direct internal argument of unaccusatives receives a theta-role from the predicate but does not receive Case, either from the predicate or from a preposition. Since the Case Filter disallows non-case-marked lexical NPs, the argument must move into a subject position, where it can receive Case. Another mechanism that ensures that unaccusatives (like passives) get subjects is the Extended Projection Principle, which requires all verbs to have subjects. The fact that the internal argument of unaccusatives is caseless is predictable because these verbs do not assign a theta-role to an external argument. This principle has become well-known under the name of Burzio's Generalization:

(24) Burzio's Generalization:
If a verb does not assign a theta-role to its subject, it does not assign Case to its object

This principle holds for unaccusatives, passives and subject raising verbs: in all these cases, the verb does not theta-mark a subject, systematically lacks accusative case , and is superficially intransitive. Less obviously, the principle also holds for unergative verbs, which can be regarded as potential accusative assigners that happen not to be categorized for a direct object.[4]

3.4.1 Diagnostics of unaccusativity

As we saw earlier, unaccusative verbs are characterized by a number of syntactic properties which distinguish them from unergative verbs. The rest of this section will focus on two of these properties, which are the most relevant for the purposes of the present study:

(a) systematic possibility of NE-cliticization, i.e. of pronominalizing a quantified i-subject NP in the form of the clitic pronoun NE, stranding a quantifier element;
(b) systematic selection of perfective auxiliary ESSERE.[5]

3.4.1.1 NE-cliticization

It appears that the distribution of NE is entirely predictable in certain syntactic configurations for inverted subjects: systematically possible with unaccusative verbs, systematically impossible with unergative verbs. As an illustration, consider the following contrasts:

(25) i. a. Sono aumentati molti prezzi
Have increased many prices
' Many prices have increased'

b. Ne sono aumentati molti
Of-them have increased many
'Many of them have increased'

ii. a. Sono venuti due amici
Have come two friends
'Two friends came'

b. Ne sono venuti due
Of-them have come two
'Two of them came'

iii. a. Parleranno due studenti
Will speak two students
'Two students will speak'

b. *Ne parleranno due
Of-them will speak two
'Two of them will speak'

It can be noted - at a descriptive level - that outside the domain of inverted subjects (i-subjects), NE-cliticization is also systematic:

(26) NE-cliticization is possible with respect to all and only direct objects

This is exemplified by the sentences below, where NE-cliticization originates from a direct object (a), an indirect object (b), and a subject (c):

(27) a. Il nemico ne ha affondate molte (direct object)
'The enemy has sunk many of them'

b. *Maria ne parlerà a due (indirect object)
Maria has spoken to two of them

c. *Tre ne arriveranno (subject)
Three of them will arrive

Moreover, even within the domain of i-subjects, the distribution of NE is predictable over certain subdomains:

(a) transitive verbs: NE-cliticization is always impossible

(28) Molti studenti leggeranno questo libro
'Many students will read this book'

*Ne leggeranno questo libro molti
Of-them will read this book many

(b) passive constructions: NE-cliticization is always possible

(29) Saranno invitati quattro studenti
Will be invited four students
'Four students will be invited'

Ne saranno invitati quattro
Of-them will be invited four
'Four of them will be invited'

(c) impersonal SI: NE-cliticization is always possible

(30) Si leggeranno volentieri alcuni testi
SI will read willingly some texts
'One will read some texts willingly'

Se ne leggeranno volentieri alcuni
SI of-them will read willingly some
'One will read some of them willingly'

In both cases (29) and (30) - passives and impersonal SI-constructions - the i-subject is clearly related to a direct object, as it is in the case of unaccusative verbs. In both cases, S-structure subjects are D-structure direct objects. The only difference is that the i-subject of passives is a direct object at D-structure level, but not at S-structure level: if the generalization in (26) is taken to refer only to D-structure direct objects, then it would be false, given cases like (31) where the NP affected by NE-cliticization is a direct object in D-structure but not in S-structure:

(31) *Molti ne saranno invitati
Many of-them will be invited

Accordingly, Burzio claims that the S-structure position must be relevant, and that the generalization in (26) refers to S-structure direct objects. The assumption therefore is that the NP in (29) has never been moved from D-structure position, so that passives fall under (26) as is. The same applies to the NP in (30), exemplifying the impersonal SI-construction. Finally, unaccusative verbs also allow the identification of i-subjects with direct object position. The case in (25a) above has the following D-structure:

(32) [e] aumentare i prezzi
increase the prices

This D-structure would give rise to cases like (33b) when NP-movement applies, and to (33a) when it fails to apply, as we saw earlier:

(33) a. I prezzi sono aumentati
prices have increased

b. Sono aumentati i prezzi
prices have increased

The generalization in (26) therefore covers unaccusative verbs as well, without further qualifications. The systematic possibility of NE-cliticization from i-subjects of passives, impersonal SI-constructions, and unaccusative verbs follows from the fact that in all these cases the i-subject is a direct object because movement rules have not been applied.

A further consequence of this analysis is that i-subjects that arise from movement are not 'direct objects' in the sense relevant for (26). Two theoretical suggestions have been made which allow this result. One - adopted by Belletti and Rizzi (1981) and dealt with in more detail below - consists of assuming that such i-subjects are not sisters of V but rather adjoined to VP, and that the syntax of NE discriminates between these two positions. The other solution, which is the one adopted by Burzio, assumes that NE-cliticization applies only to those NPs that are direct objects at all levels. This would exclude i-subjects of transitive and unergative verbs, regardless of their position in S-structure. Essentially, the two accounts differ on the basic status of NE as a clitic: while Burzio assumes that NE is base-generated, Belletti and Rizzi adopt a movement analysis of NE.[6]
In order to compare the two alternatives, the following section will present Belletti and Rizzi's proposal in some detail.

3.4.1.1.1 The syntax of NE: Belletti and Rizzi's account

3.4.1.1.1.1 Basic claims

Belletti and Rizzi (1981) start with a short descriptive account of the behaviour of NE-cliticization.

Given an indefinite quantifier NP of the form $[_{NP}Q\ N]$, the N modified by the quantifier can be pronominalized in two different ways: either with clitic NE or with a phonetically null pronoun Ø. The occurrence of these two options obeys systematic distributional constraints:

(a) in preverbal subject position, only Ø is possible:

(34) a. Due mesi passano rapidamente
Two months elapse rapidly

b. Due Ø passano rapidamente
Two Ø elapse rapidly

c. *Due ne passano rapidamente
Two ne (of them) elapse rapidly

(b) in object position, only NE is possible

(35) a. Paola trascorrerà due mesi a Roma
Paola will spend two months in Rome

b. *Paola Ø trascorrerà due a Roma
Paola will spend two in Rome

c. Paola ne trascorrerà due a Roma
Paola will spend two of them in Rome

(c) in (VP) adverbial NP's, both pronominal options are excluded

(36) a. Paola è rimasta due mesi a Roma
Paola stayed two months in Rome

b. *Paola è rimasta due Ø a Roma
Paola stayed two in Rome

c. *Paola ne è rimasta due a Roma
Paola stayed two of them in Rome

(d) under subject inversion, there are two distinct cases:

(i) postverbal subjects with verbs taking ESSERE as a perfective auxiliary behave like objects:

(37) a. Sono passati due mesi
Two months have elapsed

b. *Sono passati due
Two have elapsed

c. Ne sono passati due
Two of them have elapsed

(ii) postverbal subjects with verbs taking AVERE as a perfective auxiliary behave like adverbial NP's; i.e., both pronominal options are excluded:

(38) a. Hanno parlato due studenti
Two students have spoken

b. *Hanno parlato due
Two have spoken

c. *Ne hanno parlato due
Two of them have spoken

The two relevant phenomena here are (1) the preverbal subject/object asymmetry, and (2) the postverbal subjects asymmetry (or AVERE/ESSERE asymmetry).

3.4.1.1.1.2 Preverbal subject/object asymmetry

Belletti and Rizzi analyse the following two facts separately:

- NE pronominalization is possible only in object position
- Ø pronominalization is possible only in preverbal subject position.

As to the former, they assume a movement analysis of cliticization (Kayne 1975) and a cliticization rule

(39) adjoin clitic to V

Such a rule generates both the grammatical (b) and the ungrammatical (a) sentences in (40):

(40) a. *Due ne passano rapidamente
Two of them have elapsed quickly

b. Paola ne trascorrerà due a Roma
Paola will spend two of them in Rome

Sentence (40a) , however, is ruled out by a 'proper binding condition', a well-formedness constraint on antecedent-trace relations requiring that traces be c-commanded at S-structure by their antecedents. In (40a). the clitic does not c-

command its trace, therefore the sentence is ruled out by the proper binding condition. On the other hand, the condition is satisfied in (40b), which is well-formed.[7]

As to the latter question (i.e. why Ø-pronominalization is possible only in preverbal subject position), the solution proposed by Belletti and Rizzi relies on the assumption that the Ø pronoun is to be assimilated to PRO, the phonetically null element found in control structures. This element has a peculiar status within the GB framework because of its dual nature with respect to binding theory. PRO can be assimilated to pronominal elements (sharing the properties of ordinary definite pronouns) on the one hand, and to anaphors on the other (sharing the properties of reciprocals, reflexives, NP-traces, etc.). Such a dual nature has the effect of placing heavy constraints on the occurrence of PRO: as a pronominal, it should be free in its governing domain, whereas as an anaphor, it should be bound in its governing category. These contradictory requirements can only be satisfied when PRO is not assigned a governing category, because only in this case are both principles (vacuously) satisfied. It follows that PRO must be ungoverned, since if it had a governing category, either one or the other of the binding requirements above would be violated.

Turning to the asymmetry concerning Ø pronominalization, Belletti and Rizzi argue that PRO is assigned a governing category in the object position, but not in the subject position. Their argument departs slightly from Chomsky's (1981) original definition of government, in that in Chomsky's theory both VP-internal NP positions and the subject position in tensed clauses are governed, the former by the verb, the latter by the inflectional specification (as in the fundamental base rule S ----> NP INFL VP). Belletti and Rizzi reason that "...the inflectional element is not (necessarily) a governer, at least in the sense which is relevant for the theory of binding" (p.121). This in turn leads them to conclude that (a) the subject position is an ungoverned position in S, and (b) the object position is a governed position (governed by V).[8]

The argument, however, allows for the possibility of having sentences like (34b) above (where PRO is in an ungoverned subject position) but does not directly rule out sentences like (35b). This can be easily seen if we consider that the configurational structure of (34b) is $[_{S'}[_{S}[_{NP}$ due PRO$][_{VP}$ passano$]]]$, while the structure of (35b) is $[_{S'}[_{S}[_{NP}$ Paola$][_{VP}$ trascorrerà$][_{NP}$ due PRO] a Milano$]]]$.

In order exclude the latter sentence, it is necessary to assume that PRO is governed by V. Although at first glance this hypothesis is incompatible with the requirement that maximal projection boundaries, such as NP and S boundaries, are barriers for government, Belletti and Rizzi show that an extended definition of government is able to rule out (35b): such extended definition allows the government relation to hold between the verb and the lower projection by making the NP boundary 'transparent'.[9]

3.4.1.1.1.3 The AVERE/ESSERE asymmetry

In Italian a subject NP can generally appear in postverbal position. The crucial property to be explained here concerns the pronominal options available for postverbal subjects, which appear to be dependent on the choice of the auxiliary (AVERE or ESSERE):

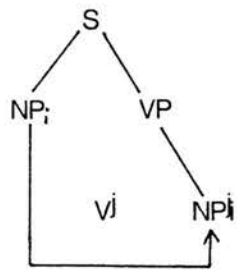
(a) the PRO option is systematically excluded in postverbal subjects
(b) the NE option is allowed when the auxiliary required by the verb is ESSERE and excluded when the auxiliary is AVERE.

The first question is related to the preverbal subject/object asymmetry discussed above. The postverbal subject is VP-internal and therefore governed by the verb: this rules out the PRO option in the same way as for the object NP's.

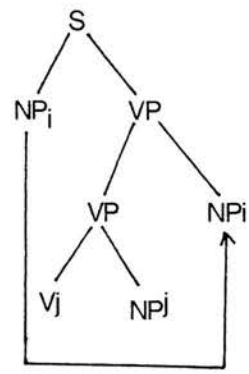
Belletti and Rizzi's treatment of the AVERE/ESSERE asymmetry is interesting because their initial assumptions are different from Burzio's, specifically in that they do not distinguish a class of unaccusative verbs. The two accounts are nevertheless compatible, as the authors explicitly recognize (Belletti and Rizzi 1981: 150; Burzio: 74).

Crucial to Belletti and Rizzi's argument is the existence of a rule "Move-NP-to-the-right" for subject inversion. This rule is able to adjoin a subject NP to both VP and S, performing either a substitution or an adjunction operation, respectively.

(41) a. Substitution



b. Adjunction



According to Belletti and Rizzi's analysis, the ESSERE case allows both outputs of the rightward NP movement rule, whereas the AVERE case only allows the adjunction output.[10] As mentioned above, the distribution of NE discriminates between the two cases: subjacency blocks NE-cliticization from an adjoined i-subject (although a slightly modified definition of subjacency is required, which will not be examined here: see Belletti and Rizzi 1981: 129).

The problem therefore reduces, in Belletti and Rizzi's view, to explaining why there can only be an adjunction (and never a substitution) output in the AVERE case. The explanation for the impossibility of the substitution output of Move-NP-to-the-right in the AVERE case lies, according to Belletti and Rizzi, in the interaction between Case theory and the theory of thematic relations. The two relevant conditions are reported below:

(42) Case Filter
Every phonetically specified NP will be Case marked

(43) Theta-criterion
a. Each contentive element (argument) fulfills one and only one theta-role
b. each theta-role is fulfilled by one and only one contentive element (where contentive elements are: lexically specified NPs, pronominals, lexical anaphors, variables and clauses).

Let us consider subject inversion with presentational *there*, as in (44),

(44) There came a man

This is a situation in which the postverbal subjects as it were 'inherits' both Case and a thematic relation from the position filled by the 'dummy' element.[11] Italian structures with subject inversion correspond to (44) in that the preverbal position vacated by the application of Move-NP-to-the-right acts as a dummy, thus transmitting Case and thematic relations to the inverted subject.

The choice of the aspectual auxiliary correlates with the Case marking property of a verb: all and only verbs that take AVERE are accusative Case assigners, as Burzio's Generalization in (24) above states. This is obvious for transitive verbs; although less obvious, it also applies to those intransitive verbs that take AVERE, which generally admit an "internal object" (e.g. *sognare un sogno* 'dream a dream', *parlare una lingua* 'speak a language') - a possibility that is never available for verbs which take ESSERE. Belletti and Rizzi therefore conclude that only intransitives taking ESSERE are 'real intransitives', which never assign objective Case, nor a thematic role to objects. Intransitives taking AVERE, on the other hand, can be considered in all relevant respects transitive verbs (i.e. objective Case assigners and assigners of theta-role to objects).

It follows from these premises that the substitution output of Move-NP-to-the-right with verbs taking AVERE is ruled out because it leads to a Case conflict, since the postverbal subject is in a context of objective Case assignment but at the same time receives nominative Case from the vacated preverbal subject position. The conflict cannot be solved by assuming that the dummy does not transmit nominative Case to the postverbal subject: this would violate the Theta-criterion, since the dummy is not a 'contentive element' and is therefore not expected to fulfill a thematic role.[12]

It is interesting to compare Belletti and Rizzi's account of the AVERE/ESSERE asymmetry with Burzio's. By assuming the existence of a natural class of unaccusative verbs, Burzio reaches very different conclusions as to which verbs are 'real intransitives' (i.e. unergative): within his theory, real intransitives are those verbs which take AVERE, because those which take ESSERE are derived from a transitive structure. Both Belletti and Rizzi and Burzio assume that NE-cliticization is possible only from i-subjects that are sisters of V like direct objects. But while for Belletti and Rizzi this configuration is the result of NP-movement, for Burzio such i-subjects are direct objects, because movement does not apply.

3.4.1.2 ESSERE assignment

Let us now turn to Burzio's account of the mechanisms of auxiliary selection.

The general distribution of the two perfective auxiliaries ESSERE and AVERE is a reliable diagnostic to distinguish unaccusative from unergative verbs. Once again, one finds regularities over well-defined syntactic domains:

(a) transitive verbs always take AVERE;
(b) passive and impersonal SI-constructions always take ESSERE;
(c) the second member of AVB/BV pairs always take ESSERE, so that within such pairs one systematically finds the auxiliary alternation in (45):

(45) a. Il governo ha aumentato i prezzi
'The government has increased the prices'

b. I prezzi sono aumentati
'The prices have increased'

Before recapitulating the above facts within specific formal generalizations regarding auxiliary selection, it is worth noticing the overlap of the distribution of auxiliary ESSERE with past participle (pp) agreement. One finds both in the following cases: passives, cases of reflexive SI, and unaccusative verbs. In such cases, the pp agrees in gender and number with the subject.

(46) a. PASSIVE: Paola è stata invitata (ESSERE; pp agreement)
'Paola has been invited'
b. REFLEXIVE SI: Paola si è lavata (ESSERE; pp agreement)
'Paola has washed herself
c. UNACCUSATIVE V: Paola è arrivata (ESSERE; pp agreement)
'Paola has arrived'

Only in two cases are auxiliary ESSERE and pp agreement dissociated:

(a) non-reflexive direct object clitics: pp agreement, no ESSERE

(47) Lo studente la ha invitata (AVERE; pp agreement)
'The student has invited her'

(b) one variant of the SI-construction: ESSERE, no pp agreement

(48) Si è parlato con gli studenti (ESSERE; no pp agreement)
'One (pl) has spoken to the students'

However, if unergative *parlare* in (48) is replaced by an unaccusative verb, then one finds pp agreement as well (where plural agreement is assumed with SI)

(49) Si è arrivati in ritardo (ESSERE; pp agreement)
'One (pl.) has arrived late'

Burzio proposes the following generalization, which relies on the idea that "in all the cases requiring ESSERE the subject enters into a certain relation with another element, while in all the cases requiring pp agreement, it is the direct object that enters into a certain type of relation" (Burzio 1986: 55).

(a) ESSERE-assignment obtains whenever there is a binding relation between the subject and a nominal contiguous to the verb. The latter can be either a clitic or a direct object, as in the following configurations:[13]

(50) ESSERE-assignment

i. NP cl-V...

ii. NP V NP...

(b) the past participle will agree with an element holding a binding relation with its direct object. Such an element can be either a clitic or a subject, as in the following configurations:

(51) Past participle agreement

i. ...cl V NP
   └──┘

ii. ...NP V NP
   └──────┘

ESSERE-assignment and past participle agreement are therefore symmetrical systems, as can be seen in the following examples to which the syntactic analyses above have been applied:

(52) a. Paola è stata invitata t
   └──────────────────┘
   'Paola has been invited'

b. Paola si è lavata [e]
   └──┘└──────┘
   'Paola washed herself'

c. [Paola] è arrivata t
   └──────────┘
   'Paola has arrived'

The passive in (52a) and the unaccusative in (52c) are both instantiations of (50ii)-(51ii). In (52b), the relation between the clitic and the empty category triggers pp agreement, as in (50I), whereas the relation between the reflexive clitic and its antecedent triggers ESSERE assignment, as in (51i).
It is worth mentioning at this point that the generalizations expressed in (50), (51) can be extended to different types of the SI-construction which goes traditionally under the name of reflexive. Burzio distinguishes among reflexive, unaccusative and inherent-reflexive verbs, all of which involve the same clitic element SI (which agrees with the subject) and require the auxiliary ESSERE, but have different syntactic structures. Examples are given below:

(53) a. REFLEXIVE: Maria si guarda
   'Maria watches herself'

b. UNACCUSATIVE: Il vetro si rompe
   'The glass breaks'

c. INHERENT-REFLEXIVE: Giovanni si sbaglia
   'Giovanni is mistaken'

In the case of (53a), SI is a reflexive object clitic which can alternate with an object NP (*Maria guarda se stessa/Giovanni*). In (53b), SI does not have reflexive meaning and does not alternate with an object: in a sense, however, it alternates with a subject (*Giovanni rompe il vetro*). It can therefore be assumed that verbs like *rompersi* (and *capovolgersi, allargarsi, muoversi, sporcarsi, rovesciarsi...*) are unaccusative just like the AVB/BV pairs described above: in fact, they all pass the NE-cliticization test. Verbs such as (53c) do not alternate either with a subject or with an object: unlike *rompersi*, they are not transitive structures. Like *rompersi*, however, they pass the unaccusativity test. They may therefore also be considered as unaccusative verbs which exhibit the affix SI and happen to lack a transitive counterpart.

The hypothesis that verbs taking inherent-reflexive and unaccusative SI should be analysed as unaccusative verbs receives further confirmation from the systematic requirement that they take both ESSERE and pp agreement:

(54) a. [La tazza] si è rotta t (ESSERE; pp agreement)
'The cup broke'

b. [Maria] si è sbagliata t (ESSERE; pp agreement)
'Maria was mistaken'

Different accounts are therefore provided for ESSERE and pp agreement with reflexive SI on the one hand, and for unaccusative and inherent-reflexive SI, on the other. The reasons for this difference will not be considered in detail, since they fall outside the scope of the present review (see Burzio 1986, chapter 1: 58-62).

Finally, the rules expressed in (50)-(51) provide a principled account of the two cases of dissociation between ESSERE and pp agreement (47-48 above, reproduced below as (55a,b):

(55) a. Lo studente la ha invitata [e]
'The student has invited her'

b. [e] si è parlato con gli studenti
'One (pl) has spoken to the students'

The two cases in (55) are instantiations of the configurations in (51ii) and (50i) respectively, which correctly predict the occurrence of pp agreement only in (55a) and of ESSERE assignment only in (55b).

3.4.1.3 Restructuring rules in Italian

The Government-Binding analysis has shed further light on other aspects of auxiliary selection that are syntactically related to unaccusativity (Rizzi, 1982; Burzio, 1986; Pearce, 1990). Rizzi (1978), in his seminal paper "A Restructuring rule in Italian Syntax" first noted that certain Italian verbs taking infinitival complements present a number of exceptional syntactic properties.[14] The classes of verbs identified by Rizzi are modals (*volere* 'want', *potere* 'can', *dovere* 'have to'), aspectuals (*cominciare* 'begin', *continuare* 'continue', *stare per* 'be about to'), and motion verbs (*andare* 'go', *venire* 'come', *tornare* 'come back'). The exceptional behaviour of these verbs is illustrated below:

(a) Clitic climbing (Cl-Cl): an unstressed pronoun originating in the embedded verb may cliticize either to the main or to the embedded verb, whereas with other main verbs only the second type of cliticization is allowed:

(56) a. Paola lo vuole comprare
b. Paola vuole comprarlo
'Paola wants to buy it'

c. *Paola lo desidera comprare
d. Paola desidera comprarlo
'Paola wishes to buy it'

(b) Long object preposing: in impersonal SI constructions, the direct object of the embedded verb may become the subject of the main verb; with other main verbs this promotion is impossible:

(57) a. Si comincia a costruire le nuove case
b. Le nuove case si cominciano a costruire
'One (pl) begins to build the new houses'

c. Si desidera costruire le nuove case
d. *Le nuove case si desiderano costruire
'One (pl) would like to build the new houses'

(c) Change of auxiliary (CA): when the embedded verb selects ESSERE, this auxiliary can be 'transmitted" to some main verbs that take AVERE; with other main verbs, this process is excluded:

(58) a. Paola ha voluto tornare a casa
b. Paola è voluta tornare a casa
'Paola wanted to go home'

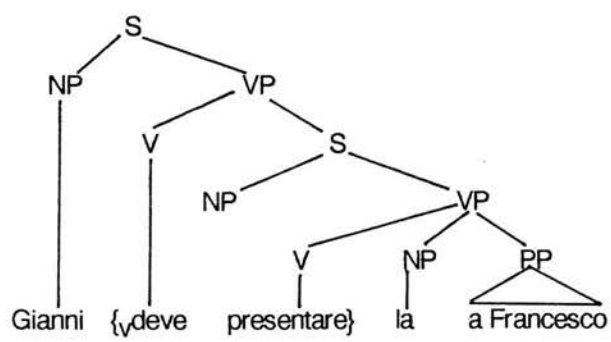
c. Paola ha promesso di tornare a casa
d. *Paola è promessa di tornare a casa
'Paola promised to go home'

3.4.1.3.1 Restructured sequences as single verbs: Rizzi's solution

As Rizzi convincingly shows, these three phenomena are interrelated, since:
(a) they have identical distribution;
(b) they react identically to certain syntactic diagnostics (see below);
(c) they tend to "go together", in the sense that if one of them occurs, the others do too.

Rizzi argues for the existence of a restructuring rule in Italian syntax, governed by modal, aspectual and motion verbs, which optionally reanalyses a terminal substring V_X (P)V as a single verbal complex, therefore transforming the underlying bisentential structure into a simple sentence (cf. Rizzi 1982: 5). The example he provides is the following:

(59)



If restructuring applies to (59), resulting in (60a), nothing will prevent the clitic pronoun from moving to the main verb, yielding (60b):

(60) a. Gianni [$_V$deve presentare] la a Francesco
b. Gianni la deve presentare a Francesco
'Gianni has to introduce her to Francesco'

If restructuring does not apply to (59), then the structure remains bisentential as in (61a) and clitic climbing on (61b) is forbidden by the SSC (Specified Subject Condition):[15]

(61) a. Gianni deve [$_S$presentare la a Francesco]
b. Gianni deve presentarla a Francesco
'Gianni has to introduce her to Francesco'

A number of syntactic derivations involving the movement of sentential or NP constituents are blocked by the presence of restructuring, because, according to Rizzi's account, such structures cease to be constituents as a effect of restructuring:[16]

(a) Wh-movement with pied-piping of the infinitival: restructuring cannot apply when the whole embedded clause is wh-preposed, as in (62b); it can when only the PP immediately containing the wh-word is preposed, as in (62c).

(62) a. Maria avrebbe voluto tornare alla casa paterna
b. *La casa paterna, tornare alla quale Maria sarebbe voluta...
c. La casa paterna, alla quale Maria sarebbe voluta tornare...
'Maria would have liked to go back to her parents' house'

(b) Clefting: restructuring is incompatible with clefting of the infinitival complement, as illustrated by the impossibility of Change of Auxiliary in 63b) and of Clitic-Climbing in (63d)

(63) a. E' andare a casa che Paola avrebbe dovuto
b. *E' andare a casa che Paola sarebbe dovuta
'It is going home that Paola was supposed to do'

c. E' proprio comprarlo che Paola voleva
d. *E' proprio comprare che Paola lo voleva
'It is actually buying it that Paola wanted'

(c) Right Node Raising: restructuring is incompatible with this rule operating on coordinate sentences whose rightmost constituents are identical. This rules out (64c) and (64f):

(64) a. Paola vorrebbe pagargli il suo debito, ma non potrà mai pagargli il suo debito
'Paola would like to pay him her debt, but will never be able to pay him her debt'

b. Paola vorrebbe - ma non potrà mai - pagargli il suo debito
'Paola would like - but will never be able to - pay him her debt'

c. *Paola gli vorrebbe - ma non gli potrà mai - pagare il suo debito

d. Maria avrebbe voluto/sarebbe voluta venire alla festa, ma non ha potuto/è potuta venire alla festa
'Maria would have liked to come to the party, but could not come to the party'

e. Maria avrebbe voluto - ma non ha potuto - venire alla festa
f. *Maria sarebbe voluta - ma non è potuta - venire alla festa
'Maria would have liked - but could not - come to the party'

(d) Complex NP shift: again, restructuring appears to be incompatible with this rule which shifts a complex postverbal complement in sentence-final position, as the ungrammatical Change of Auxiliary in (65d) shows:

(65) a. Il fiume ha cominciato a straripare nelle campagne vicino a Pisa
b. Il fiume è cominciato a straripare nelle campagne vicino a Pisa
c. Il fiume ha cominciato vicino a Pisa a straripare nelle campagne
d. *Il fiume è cominciato vicino a Pisa a straripare nelle campagne
'The river began to overflow in the country near Pisa'

The interaction among the three phenomena is particularly evident in the case of Clitic-Climbing and Change of Auxiliary. As (66) shows, when Clitic-Climbing occurs, Change of Auxiliary cannot fail (although the reverse does not hold):[17]

(66)	a.	Maria avrebbe voluto andarci	(no Cl-Cl; no CA)
	b.	Maria sarebbe voluta andarci	(no Cl-Cl; CA)
	c.	Maria ci sarebbe voluta andare	(Cl-Cl; CA)
	d.	*Maria ci avrebbe voluto andare	(Cl-Cl; no CA)
		'Maria would have liked to go there'	

Rizzi's formulation of restructuring rules to account for the AVERE-->ESSERE change is the following:

(67) AVERE --> ESSERE / [$_V$vbl_____vbl V_k]

The rule applies to occurrences of AVERE only where V_k is a verb requiring ESSERE.[18] This rule is not entirely unsatisfactory, as Rizzi himself recognizes: "It would be highly desirable not to have a specific rule at all for these cases, with the paradigm discussed...being predicted by some general principle for auxiliary assignment interacting with Restructuring.....but only a more detailed study of the syntax of auxiliaries could permit a satisfactory account of this phenomenon" (p. 23).

3.4.1.3.2 Restructuring as VP-movement: Burzio's solution

Burzio (1986: chapter 5) attempts to incorporate Rizzi's results and at the same time to go beyond them, by providing an explanatory account of auxiliaries in restructured constructions. Such an account is made possible by assuming the existence of the class of unaccusative verbs, which Rizzi failed to recognize.

Burzio's idea is basically to treat restructuring as syntactically derived and not base-generated.[19] In his account, the restructuring rule as an instance of VP-movement: it consists of a permutation of the linear order of direct object and infinitive, such that the infinitival VP moves to the left, out of its clause. Structurally, this permutation can be represented as follows (where V_1 is an unaccusative verb, NP_1 its direct object, and S its complement):

(68) a. $[_{VP1}\ V_1\ NP_1\ [_S\ PRO[_{VP2}\ \mathbf{V_2\ NP_2}]]]$ non-restructured

b. $[_{VP1}\ V_1\ [_{VP2}\ \mathbf{V_2\ NP_2}]\ NP_1[_S\ PRO\text{- - -}]]$ restructured

Notice that the embedded subject is not deleted, although it is never phonologically realized.[20] We shall return to this point later.

The effects of the restructuring process are illustrated in (69), (70), and (71) for each of the main types of inducing verbs:

(69) a. Unaccusative: *andare, venire*

b. $Paola_i$ va t_i $[_S PRO_i$ a comprare il latte] --->
$Paola_i$ va $[_{VP}$ a comprare il latte] t_i $[_S\ PRO_i$---]
'Paola goes to buy the milk'

(70) a. Raising: *dovere, potere, cominciare, continuare, stare per, sembrare*

b. $Paola_i$ dovrebbe $[_S\ t_i$ comprare il latte] ---->
$Paola_i$ dovrebbe $[_{VP}$ comprare il latte] $[_S\ t_i$---]
'Paola should buy the milk'

(71) a. Control: *volere, sapere, cominciare, continuare*

b. $Paola_i$ vorrebbe $[_S\ PRO_i$ comprare il latte]--->
$Paola_i$ vorrebbe $[_{VP}$ comprare il latte] $[_S\ PRO_i$---]
'Paola would like to buy the milk'

What follows will be essentially concerned with auxiliary change under restructuring, which is also the main focus of Burzio's account of restructuring; Burzio admits his 'incomplete' understanding of clitic-climbing[21] (for the many other aspects of restructuring rules that will not be covered in this chapter, see Burzio 1986, chapter 5).

Consider the alternation in (72) below, which exemplifies the Change of Auxiliary phenomenon:

(72) a. Paola$_i$ avrebbe voluto [$_S$ PRO$_i$ venire t$_i$]

b. Paola$_i$ sarebbe voluta [$_{VP}$ venire t$_i$] [$_S$ PRO$_i$---]

'Paola would have liked to come'

Notice first that this construction can be regarded configurationally as an extension of the core configuration for ESSERE-assignment in (50i) above, i.e. NP V NP. If the notion of direct object is based on government, rather than on thematic features, then the subject of infinitives can in certain cases behave just like the direct object. This explains why verbs such as *sembrare* appear with both auxiliaries, like unaccusative verbs. What happens in (72b) is that as a result of VP-movement the direct object of the embedded verb (a trace) enters in a direct relation with the matrix subject, which is independently coindexed with it and which c-commands it. Since in complex predicates the object of the embedded verb functions as the object of the matrix verb as well, the new relation induces ESSERE on the main verb. Clearly such an account is made possible only by an unaccusative analysis of the verb *venire*.

The inadequacy of Rizzi's rule for Change of Auxiliary in (67) is even more apparent if one considers its inability to explain the directional asymmetry of the change, i.e. the fact that Change of Auxiliary occurs only (a) from AVERE to ESSERE and (b) from right to left.[22] An example of the former constraint is (73):

(73) a. Maria$_i$ è andata$_i$ t$_i$ [$_S$PRO$_i$ a comprare il latte] (non restructured)

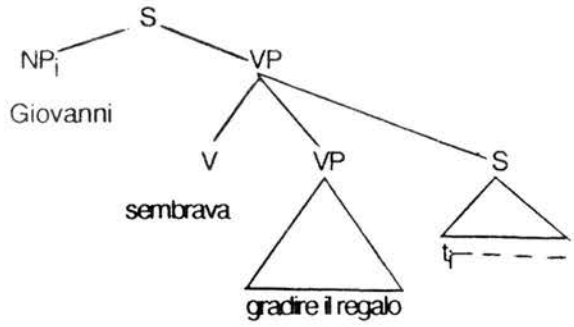
b. Maria$_i$ è andata [$_{VP}$ a comprare il latte] t$_i$ [$_S$ PRO$_i$---] (restructured)

'Maria went to buy the milk'

In these cases of restructuring with *andare*, the matrix verb continues to take ESSERE after the application of the rule, i.e. there is no transmission of auxiliary AVERE from the embedded transitive verb *comprare*. This is due to the continuing presence of the direct object t$_i$ linked with the subject: all structure is preserved under restructuring.

The directionality of Auxiliary Change can be explained by assuming that the complex predicates in restructuring constructions are structurally asymmetric, with the leftmost verb higher in the structure than the rightmost and c-commanding the latter, as in (74):

(74)



The object of the main verb t_i cannot become the object of the embedded verb, since the latter does not c-command that object. The main verb cannot therefore transmit auxiliary ESSERE to the embedded one: there is not left-to-right Change of Auxiliary, as (75) shows (the position of the clitic is evidence that restructuring has occurred):

(75) a. Giovanni lo era sembrato gradire
Giovanni had seemed to like it

b. Giovanni lo sembrava aver gradito
Giovanni seemed to have liked it

Finally, the VP-movement analysis of restructuring is further confirmed by the distribution of auxiliaries over Raising configurations.
In the absence of restructuring, Raising can be represented as follows (NP_2 is the trace of NP_1:

(76) $NP_1\ V_1\ [_S NP_2\ V_2\ (NP_3)...]$

Consider now non restructured Raising (77a) and restructured Raising (77b):

(77) a. Paola$_i$ ha potuto [$_S$ t_i venire t'$_i$]

b. Paola$_i$ è potuta [$_{VP}$ venire t'$_i$] [$_S$ t_i---]
'Paola could come'

Auxiliary AVERE of *potere* exemplifies a case in which the relation between NP_1 and NP_2 does not count for ESSERE assignment (see section 3.5 for an explanation). The Change of Auxiliary in (77b) indicates that the direct object of the embedded

verb, NP_3 of (76), is reanalysed, as a result of restructuring, as the object of the main verb. NP_2, however, is not so reanalysed. This is shown by (78) (where we ignore the fact that the clitic, whose only function is to signal the presence of restructuring, has an empty category in the embedded VP):

(78) Paola$_i$ mi ha potuto [$_{VP}$ parlare] [$_S$ t$_i$---]
Paola could speak to me

If the trace in (78) - which corresponds to NP_2 in (76) - were reanalysed as a direct object of the main verb, then *potere* should take ESSERE even though the embedded verb is not unaccusative. NP_2, however, is not deleted either by restructuring, as shown by the fact that other peripheral cases, like (75a) above, select ESSERE.

The consequences of restructuring on Raising can therefore be summarized as follows:

(a) the relation between V_i and NP_2 remains unchanged (NP_2 is not deleted);
(b) NP_3 is reanalysed as an object of V_i.

This conclusion, Burzio argues, excludes any view of restructuring that does not involve movement: if no movement occurred, NP_2 would be reanalysed before NP_3 in (76).

To sum up, the GB analysis of unaccusativity makes it possible to account for both the selection of perfective auxiliaries and optional auxiliary change under restructuring on the basis of the same set of theoretical premises. The restructuring phenomena remain completely unexplained within semantic theories: the derivation for these phenomena therefore appears to be exclusively syntactic.

3.5 The French auxiliary system

French has roughly the same class of unaccusative verbs as Italian and shares a number of syntactic regularities with it: transitive and unergative verbs take AVOIR, while ETRE is found with reflexive and passive constructions

(79) a. Marie a mangé une pomme — transitive
'Marie has eaten an apple'

b. Ma soeur a voyagé en train — unergative
'My sister has travelled by train'

c. La loi est soutenue par le gouvernement — passive
'The law is supported by the Government'

d. Paul s'est rasé — reflexive
'Paul has shaved'

The variability of French auxiliary selection with unaccusative verbs can be descriptively summarized as follows:

(a) all paired unaccusatives in reflexive form and inherently reflexive verbs take ETRE;

(b) only a subset of unpaired unaccusatives take ETRE, whereas the majority take AVOIR ;

(c) paired unaccusative in non-reflexive forms select AVOIR.

(80) a. Le vase s'est cassé — paired unaccusative, (reflexive form)
'The vase broke'

b. Paul ne s'est aperçu de rien — unpaired unaccusative, (inherently reflexive)
'Paul didn't notice anything'

c. Paul est allé à la maison — unpaired unaccusative
'Paul went home'

d. L'argent a disparu — unpaired unaccusative
'The money disappeared'

e. Tu as beaucoup changé — paired unaccusative, non-reflexive
'You changed a lot'

As stressed by the Government-Binding analysis of French auxiliaries, some of the differences stem from the fact that French - unlike Italian - is a non pro-drop language.

The phenomena associated with IL-inversion in French are a primary example of this. As a non pro-drop language, French lacks inversion by rightward movement, due to a constraint that limits insertion of pleonastic IL in D-structure (see Burzio 1986: 163). Such a constraint does not operate in Italian, because Italian allows null subjects and therefore does not require insertion (in fact the complete productivity of inversion in Italian is itself a reflex of the Null Subject property).

The second type of inversion involves non-argument IL, which appears to discriminate between certain verbs and others:

(81) a. Il est arrivé trois étudiants

b. ?Il a téléphoné trois étudiants

Moreover, IL-inversion generally allows cliticization of EN, which is just like its Italian counterpart NE. The verbs allowing both inversion and EN-cliticization are precisely the class of unaccusative verbs; conversely, the two phenomena are impossible with unergative verbs.

Now it may seem plausible to assume that the two French auxiliaries ETRE and AVOIR are exactly equivalent to Italian ESSERE and AVERE and therefore share the same system of auxiliary assignment. If this were the case, one would expect IL-inversion to be possible with all and only the verbs that take auxiliary ETRE. Half of this prediction seems to be fulfilled: French verbs that take ETRE (*arriver, aller, venir*, etc.) generally allow inversion. The other half of the prediction is not entirely fulfilled because one finds cases of IL-inversion with verbs taking AVOIR (which nevertheless allow EN):

(82) a. Il a disparu de l'argent de ma poche /Il en a disparu beaucoup
b. Il a manqué trois étudiants /Il en a manqué trois

Interestingly, however, the corresponding Italian verbs (*sparire, mancare*, etc.) take ESSERE.
Of the two possible solutions - either that French has a different class of unaccusative verbs, or that French has a different system of auxiliary assignment - the latter seems to be true. In the configuration in (50ii), shown in (83) below, which is the one relevant to unaccusative verbs, lexical factors play a role in French, though not in Italian:

(83) NP V NP
└──┘

Such lexical factors would be responsible for the fact that verbs like *arriver* take ETRE and verbs like *disparaître* take AVERE, whereas in Italian all the corresponding verbs take ESSERE. As will be seen later, this could be the result of the diachronic effect

of progressive 'erosion' of the derivations of Latin esse in the auxiliary systems of Romance languages: French is more advanced than Italian in the historical trend that favours derivations of *habere*. There is a sense in which one could say that the Italian system of auxiliary assignment - with the whole cluster of properties associated with it - has an internal syntactic coherence that the French system has lost. We will come back to this issue in section 3.9.3.

At this point it is necessary to reconsider the syntactic configurations for ESSERE assignment in Italian that were presented in (50i, ii) above and are reproduced in (84):

(84) a. NP Cl-V

b. NP V NP

It appears, in fact, that although the Italian system is more consistent than the French one, it also has an area of idiosyncrasy (though more restricted than French). This area is represented by Raising configurations like (85a), exemplified by (85b), (85c):

(85) a. NP V [$_S$NP...]

b. Paola ha dovuto accompagnare i figli a scuola
'Paola had to take her children to school'

c. Paola è sembrata avere dei problemi
'Paola seemed to have problems'

Raising configurations in Italian seem to allow either auxiliary, whereas the corresponding French sentences never allow ETRE:

(86) a. Pauline a du emmener les enfants à l'école

b. Pauline a semblé avoir des problèmes

However, French always allows ETRE in configuration (84a). The behaviour of the auxiliaries ESSERE/ETRE can therefore be summarized by assuming that

both languages have core cases for ESSERE/ETRE assignment (in which ESSERE/ETRE cannot fail) and peripheral cases (which allow ESSERE/AVERE or ETRE/AVOIR alternations). All cases that fall ouside the system are core cases for AVERE/AVOIR assignment.

The differences between Italian and French can be stated in terms of a hierarchy of core and peripheral syntactic configurations for ESSERE/ETRE assignment:

TABLE 3.1: Configurational hierarchy for ESSERE-and ETRE-selection in Italian and French (adapted from Burzio, 1986: 140)

	ITALIAN	FRENCH	
a. NP cl-V	core	core	(unaccusatives in reflexive form)
b. NP V NP	core	periphery	(unaccusatives in non-reflexive form)
c. NP V [$_S$ NP...]	periphery	--------	(Raising and restructuring constructions)

Of the three syntactic configurations inducing binding, and therefore ESSERE/ETRE-assignment, the first two are defined as core (that is, regular and systematic) in Italian and the third as peripheral (that is, allowing variation), whereas in French only the first configuration is core, the second one is peripheral, and the third one is outside the system altogether (and thus a core case for AVOIR-assignment). The configuration (b) in Table 3.1 accounts for paired and unpaired non-reflexive unaccusative verbs, and is characterized as core in Italian and as peripheral in French. In fact, this is where one finds variation and apparent unpredictability in auxiliary selection.

It will be recalled that ETRE assignment is triggered by a specific relation between the subject and an element contiguous to the verb. Burzio notes that the cases above are ordered precisely with respect to the degree of proximity of the relevant element to the verb. The rule of ESSERE/ETRE assignment is therefore parameterized with respect to the degree of contiguity it requires: in more formal terms, what is parameterized is the notion of government that enters into the system (Burzio 1986: 140).[23]

The periphery was provisionally defined as an area that allows variation in auxiliary selection. However, even the periphery presents some striking regularities: if a verb ever falls into the core of the system (either the core for ESSERE /ETREor the

core for AVERE/AVOIR assignment), then the auxiliary assigned in the core is maintained in the periphery. Turning to (85b) above, the verb *sembrare*, unlike *potere*, occurs in sentences like (87), which is a core case for ESSERE assignment (it actually falls under configuration (b) in Table 3.1, once the latter is generalized to sentential complements). *Sembrare* therefore maintains ESSERE in the peripheral case.

(87) [e] sarebbe sembrato [$_S$ che Paola avesse dei problemi]
It would have seemed that Paola had some problems

The corresponding French case in (86b) is a core case for AVOIR assignment, because it falls outside the system for ETRE assignment. Again, the same principle is at work: French *sembler* takes AVOIR as a core case and maintains it in the periphery. This principle predicts another important regularity: French unaccusative verbs that have transitive alternants (i.e. belonging to the AVB/BV type, where both AVB and BV are lexicalized) ought to take AVOIR, since the peripheral configuration for unaccusative verbs will "inherit" the auxiliary from the core configuration for AVOIR assignment. This prediction is confirmed in the vast majority of cases. Three examples are given in (88):

(88) a. Ils ont coulé le bateau
b. Le bateau a coulé

c. Vous avez changé beaucoup de choses ici
d. Beaucoup de choses ont changé ici

e. Le gouvernement a augmenté les prix
f. Les prix ont augmenté

It is not only with regard to auxiliary selection that French unaccusative verbs are considerably less regular in their syntactic behaviour than their Italian counterparts. Unlike Italian unaccusatives, which reliably satisfy syntactic diagnostics as a class, no single syntactic diagnostic can identify the whole class of French unaccusatives. As Legendre (1989) points out, the condition of unaccusativity in French is that an intransitive verb has to pass at least one unaccusativity test:[24]

" ...distinct tests single out distinct - but often overlapping - subsets of unaccusatives....In order for a given intransitive verb v to satisfy one test X, it is necessary that v be unaccusative; in order to demonstrate that v is unaccusative, it is sufficient that v satisfies one test...All the tests together cover all French unaccusatives, so that passing at least one of these tests is a necessary condition for unaccusativity...An intransitive verb is unergative if and only if it fails all unaccusativity tests."
(Legendre 1989: 97)

While verbs in reflexive form generally satisfy the usual diagnostics for unaccusativity (for example, they allow EN-cliticization and participial adjectives), non-reflexive unaccusatives - particularly paired ones- do not easily allow either.

(89) a. Il ne s'en est cassé que trois ici
There of-them is broken only three here
'Only three of them broke here'

b. *Il n'en a coulé que trois ici
There of-them has sunk only three here
'Only three of them sank here'

c. Le garçon arrivé hier était mon frère
'The boy arrived yesterday was my brother'

d. Le bateau coulé hier (*tout seul) a gagné la compétition
'The boat sunk yesterday (all by itself) won the competition'[25]

Labelle (1990), following work by Pollock (1985), also shows that paired unaccusatives in non-reflexive form (unlike reflexive unaccusatives) do not display other syntactic characteristics, such as the ability to enter impersonal constructions or to relativize their arguments when embedded under *croire* (e.g. *Le vase que tu croyais s'être brisé* vs **La personne que je croyais avoir vieilli*). She claims that this class of verbs is not unaccusative but unergative.[26]
It seems that the drastic reduction of the scope of ETRE , and the erosion of the link between ETRE and unaccusativity, have introduced unaccusative 'mismatches' (cf. Levin, 1988; Labelle, 1990; Zubizarreta, 1987; Burzio, 1986) that have altered the internal consistency of the system. We shall return to this issue in section 3.9.3.

Finally, modern French does not exhibit any restructuring phenomena, and does not therefore allow either clitic climbing or auxiliary change. As predicted by the configurational hierarchy in Table 3.1, these constructions fall within the syntactic configuration (NP V [$_S$NP...] , and are therefore outside the system of ETRE-assignment in French (whereas they are peripheral (i.e. variable) cases of ESSERE-assignment in Italian).
Both constructions, however, were definitely attested in Old French and throughout Middle French up to the 17th century (Kayne, 1980; Pearce, 1990):

(90) a. Li mareschaux n'estoit voulu venir a lui — auxiliary change
(13thC., Livre de la Conqueste)

b. mais ce sexe par nul exemple n'y est encores pu arriver — clitic-climbing
(16thC., Montaigne, Essais I)

(Examples from Pearce, 1990: 20-21)

This, together with the fact that the grammar of Old French allowed a wider distribution for ETRE (see Gougenheim, 1951 and the final section of this chapter), confirms the deep relation between these seemingly unrelated properties. It also lends support to the view that French used to be more similar to modern Italian, and that the situation one finds in modern French is a historical development of the system exhibited by Italian.

3.6 Unaccusativity in English

3.6.1 General remarks

In comparison with Italian and French, English can be said to have roughly the same class of unaccusatives, but has only a few, if any, syntactic reflexes of unaccusativity.
Modern English does not offer a choice of auxiliaries in perfect tenses, although *be* was in use until the beginning of the 19th century (Rydén and Brorström, 1987) mainly as the auxiliary of unaccusative verbs (as for example in 'he is come'). Furthermore, English has neither clitics, nor restructuring constructions. There is therefore no equivalent to the syntactic diagnostics that identify unaccusatives in Italian and French. One possible exception is expletive *there*-insertion, which is systematically disallowed with unergative verbs and allowed by some unaccusative verbs, as (91) shows:

(91) a. *There worked many men on this building site
b. There came three new students
c. There has appeared another book by Chomsky
d. There exists a constraint on grammar development

Even *there*-insertion, however, seems to be grammatical only with a semantically identifiable subset of unaccusative verbs, which includes "change of state" or presentational verbs, and excludes verbs with a transitive alternant such as *sink* in (92) (cf. Burzio, 1986; Napoli, 1988; Haegeman, 1990).[27]

(92) *There sank three ships

3.6.2 A problem for language acquisition

The fact that English unaccusatives do not have overt morphological markers that set them aside from unergatives raises the question of how the child who acquires English as a native language knows how to classify new verbs as they are learned (see Baker 1983; Rappaport 1989; Zubizarreta 1987). If the two different syntactic representations underlying the unaccusative and the unergative verb classes are allowed by Universal Grammar, then knowledge of the distinction must eventually be acquired. This knowledge would be what enables native speakers of English to accept *there*-insertion with some unpaired unaccusative verbs and to reject it with unergative verbs.

A solution to this problem is to assume that the language learner relies on semantic criteria to determine the class membership of intransitive verbs. If this is correct, it follows that the semantics of a system like English must be sufficiently consistent to allow acquisition to take place. There is considerable convergence of opinions on this view, as the following quotes show:

> "In languages with overt morphological markers of unaccusativity, the property may be grammaticalized. Since there are overt indicators of class membership, then there may be a certain amount of deviation from the semantic criteria for unaccusativity. This would not interfere with acquisition."
> (Rappaport 1989: 120)

> "...In languages which have relatively few reflexes of the unaccusative/unergative distinction semantics alone will determine their structure. Therefore the two classes of verbs should be more easily characterizable semantically in such languages."
> (Baker 1983: 35)

Semantic criteria must also guide the language learner in acquiring the knowledge that only some unpaired unaccusative verbs allow the *there*-insertion construction and, in particular, that paired unaccusatives with a transitive alternant do not. As we saw earlier, this latter type of unaccusative display an atypical syntactic behaviour in French as well: thus, the French learner too has to acquire knowledge of the special character of these verbs. The child acquiring French, however, has overt positive evidence that verbs like *couler* (see example in (88) above) are exceptional: for example, the choice of auxiliary AVOIR instead of ETRE. The child acquiring English, on the other hand, does not find any explicit indicators in the input that *sink* is an irregular unaccusative verb. The constraint that prevents *sink* from appearing in *there* insertion constructions is to be found, according to Zubizarreta (1987: 92), at the level of the sentence: the sentence must be presentational. There is no need, therefore, to treat these type of unaccusatives in English as formally different from other unaccusative verbs. In French, however, the special behaviour of verbs such as *couler* raises the question of whether they are to be considered as members of the unaccusative class. As we shall see later, there is some theoretical justification in regarding these verbs as syntactically closer to unergative verbs, and in placing the source of irregularity at the level of the mapping between semantic and syntactic representations (Zubizarreta 1987).

3.6.3 English unaccusatives: a syntactic or a semantic class?

The precise characterization of English unaccusatives is a matter of debate in the literature. According to Keyser and Roeper (1984), English unaccusative verbs are syntactically intransitive but are generated from transitive verbs by a movement rule in the lexicon: English and Italian are said to be in parametric variation for having a lexical versus a syntactic rule of unaccusative formation, respectively. A different position (Napoli, 1988) claims that English unaccusatives are intransitive both lexically and syntactically, and that they form a semantic, rather than syntactic class. The debate on the syntactic vs semantic nature of English unaccusatives revolves around the five central arguments provided by

Keyser and Roeper in support of their claim that unaccusatives are originally transitive in the lexicon and become intransitives via a movement rule:

(a) the rule of unaccusative formation ('Ergative Rule' in Keyser and Roeper's terms) is productive;

(b) the suffix *-er* cannot attach to an unaccusative verb to produce the sense of a theme argument, but only of an agent argument. Therefore, *-er* attachment must apply in the lexicon before the Ergative Rule;

(c) the trace of lexical movement in an unaccusative verb prevents the insertion of a cognate object;

(d) *there*-insertion applies only to unaccusatives (and not to other intransitive verbs) because the NP following the unaccusative verb appears where it is generated in the lexicon;

(e) the prefix *re-* can occur only with unaccusative verbs because it requires linking with an object NP, and this requirement is satisfied by the trace of the lexical movement.

Napoli brings clear and plausible counter-arguments against the view that unaccusative verbs are originally transitive in the lexicon. She argues that each of Keyser and Roeper's arguments is empirically flawed, and that the same generalizations (and their exceptions) are best explained in semantic terms.

The flavour of Napoli's argumentation will be illustrated with respect to two of the five arguments listed above: the productivity of the 'Ergative Rule' and the rule of *-er* attachment.

According to Keyser and Roeper, the demonstrated productivity of a rule is evidence for its existence. They argue that new verbs can be formed in English which exhibit the transitive/unaccusative alternation, as in (93):

(93) a. The Republicans want to Reaganize the country
b. The country refuses to Reaganize

This rule, however, is not fully productive in the sense required by syntactic rules. It is not difficult to find verbs that do not enter the transitive/unaccusative alternation:

(94) a. Mary visualized the town
b. *The town visualized

Keyser and Roeper's conclusion is that "the intransitive member of an ergative pair must be generated by rule for each new lexical entry." (p. 390). But this is equivalent to denying the assumed productivity of the lexical Ergative Rule, because the rule is regarded as an operation which applies to lexical entries in an 'ad hoc' and essentially unprincipled way. As we saw earlier, the possibility of the transitive/unaccusative alternation (Burzio's AVB/BV pairs) is not fully productive in Italian either: the variation is not accounted for in syntactic terms but is attributed to "lexical factors". This seems to indicate - as Napoli does - that the possibility of the transitive/unaccusative alternation is variable because it results from the lexical-semantic meaning of lexical entries: Napoli explicitly refers to "information which is prelexical, perhaps even prelinguistic" as the determining factor.

As to the attachment of suffix *-er*, Keyser and Roeper note that *-er* generally attaches to subjects of transitive verbs, but not to subjects of derived intransitive verbs. For example *the baker* can only have an agentive meaning, i.e. it can only refer to *the man* in *The man baked the bread* and not to *the bread* in *the bread baked*. Keyser and Roeper rely on these morphological data to demonstrate that the input to the Ergative Rule is the transitive member of the transitive/unaccusative pair , and not the intransitive member: *-er* attachment must apply in the lexicon to transitive verbs before they undergo the putative Ergative Rule. This requirement would account for the fact that unaccusative verbs generally seem to resist the attachment of suffix *-er* to yield an agentive meaning. However, exceptions to this generalization can be easily found, as Keyser and Roeper themselves admit (they quote *roaster, broiler, sinker* as potential counterexamples). For instance, an unaccusative verb like *stick* (which participates in a transitive/unaccusative alternation: *I stuck the note on the door* vs *The note stuck fast on the door*) allows *-er* attachment both in the theme and in the agent sense, as the examples given by Napoli show:

(95) a. What good little stickers these notes of yours are! (theme)
b. Okay, let's divide up the jobs. I write the notes, you stick them up. So I'm the writer and you are the sticker. (agent)

This is what one would expect if both the transitive and the intransitive senses of *stick* (and not just the transitive sense) are available to *-er* attachment.
While it is true that the majority of unaccusative verbs do not easily allow *-er* attachment in the theme sense, the explanation is to be found, according to

Napoli, in a semantic constraint that limits this attachment to subjects that are participators in an event. So, in the case of (95a) above, the subject of *stick* has an active role in the event and therefore *sticker* can have an theme meaning. But glasses, for example, cannot be regarded as active participants in their own breaking, so that *breaker* cannot take an theme meaning (*This glass is a breaker). This semantic constraint seems to operate on unpaired unaccusatives as well.Verbs such as *arrive* do not easily allow *-er* attachment, but if we try it (as in *Who is the new arriver*?) we obtain only the sense corresponding to the person who arrives (i.e. the theme meaning), not to someone who causes the arrival.

The conclusion drawn by Napoli is that English unaccusatives are intransitive at both the lexical and the syntactic levels. They 'act together' with respect to a number of syntactic phenomena but they do so not as a syntactically coherent class, but rather in terms of semantically defined subclasses. This is in complete agreement with the other analyses of unaccusativity in English mentioned in 3.6.2.

3.6.4 The semantics of English verbs of motion

The semantic bases of English unaccusativity can be illustrated by an analysis of the class of verbs of motion. These verbs are notoriously recalcitrant to a consistent semantic characterizations, and therefore pose a problem for any attempts to define the Unaccusative Hypothesis in terms of semantic universals (see section 3.8.1 below). Agentivity and themehood, as the commonly suggested thematic features of unergative and unaccusative verbs respectively, are contradicted by verbs like *arrive*, which are agentive and yet display consistent unaccusative behaviour across languages, and verbs like *swim* , which are consistently classified as unergative even though their subject is a theme, in the sense that it undergoes a change of location.

One solution to this problem is to abandon the idea that verbs of motion are a 'grammatically relevant semantically coherent class' (Rappaport 1989), and to analyse these verbs in terms of semantically defined sub-types. It is possible to classify verbs of motion on the basis of the two dimensions '+/- telic' and '+/- protagonist control' (the latter feature seems to be equivalent to 'agentivity'). If a verb is -telic, then the class membership is determined by the feature 'protagonist control': +protagonist control verbs will be unergative, whereas

-protagonist control verbs will be unaccusative. If a verb is +telic, then it is unaccusative, irrespective of the feature of protagonist control.

TABLE 3.2: Semantic parameters for English verbs of motion

+telic		UNACCUSATIVE
-telic	-protagonist control	UNACCUSATIVE
	+protagonist control	UNERGATIVE

The three classes identified by the parameters in Table 3.2 are:

(a) the *arrive*-type, which are +telic

(b) the *roll*-type, which are -telic, -protagonist control

(c) the *run*-type, which are -telic, +protagonist control.

The *arrive*-type class includes inherently directed verbs of motion, which specify the end-point or source-point of motion. These verbs denote the achievement of a change of location, and not the process leading to it, or the manner of motion. Both the *roll*-type and the *run*-type are activity verbs, which denote unbounded actions. The difference between +telic and -telic bears on a number of syntactic diagnostics of unaccusativity. One of these is the compatibility with *resultative adjuncts*, which can be predicated only of d-structure direct objects and denote a state achieved by the referent as a result of the action expressed by the verb. Transitive verbs can have resultative adjuncts predicated of their direct objects, as in (96):

(96) Bill scrubbed the counter clean

In contrast, unergative verbs cannot have resultative adjuncts predicated of their sole argument:

(97) *I worked tired

Unaccusative verbs are split on this issue. Some can have resultative adjuncts predicated of their sole argument, which is of syntactically explainable by assuming that the s-structure subject of these verbs is a d-structure direct object:

(98) The river froze solid
The vase broke into pieces

Telic verbs of motion, however, are not compatible with resultative adjuncts:

(99) *We arrived (ourselves) ragged

The incompatibility is due to the fact that since the *arrive*-type verbs already specify an achieved change of location, they are grammatically constrained from specifying an achieved state as well because a clause cannot specify more than one achievement.[28] As Rappaport suggests, the fact that verbs of directed motion do not pattern like other unaccusatives with respect to this diagnostic does not question their membership to the class of unaccusative verbs: rather, it is explained by a general grammatical-semantic constraint. This conclusion is again in complete agreement with the other accounts of unaccusativity in English that were mentioned before.

We can conclude from this brief overview of English unaccusativity that English is different from Italian and French with regard to the syntactic behaviour of unaccusative verbs, but not with regard to the semantic characterization of these verbs. This generalization will be important for the interpretation of the empirical results of the present study.

3.7 Syntactic explanations of auxiliary selection: summary

The general issue arising from all the theoretical accounts examined so far is whether a purely syntactic approach is capable of explaining the whole range of grammatical properties related to the choice of auxiliaries. GB theory captures many of the significant generalizations in this domain. Auxiliary selection and NE-cliticization are both seen as syntactic manifestations of unaccusativity, derivable from the the particular configuration assumed by unaccusative verbs at s-structure by virtue of their having a direct object in d-structure. Optional and obligatory auxiliary alternations in the domain of Restructuring are shown to be sensitive to the same syntactic principles and ultimately to depend from the syntactic characterization of unaccusativity offered in GB terms, particularly from the notion of government. Furthermore, the GB analysis proves to be extendable to other languages, thus establishing the bases for a typology of auxiliary systems.[29]

The question remains, however, as to whether semantic considerations are incompatible with syntactic principles. This seems to be Burzio's position: he argues at length that it would be pointless to establish any correspondence between the syntactic principles responsible for auxiliary selection and thematic relations, since the mechanisms of auxiliary assignment work in the same fashion regardless of the thematic relations exhibited by individual sentences. For instance, consider (100):

(100) a. Il sasso è caduto
'The stone has fallen'
b. Maria è sembrata avere dei dubbi
'Maria seemed to have some doubt'
c. Maria è voluta venire
'Maria wanted to come'

In all these sentences - Burzio argues - what induces the selection of ESSERE is the binding relation between the subject and an element governed by the verb at s-structure, regardless of the thematic roles borne by the subject (which are different in the three cases: a patient in (a), an agent but with respect to the embedded verb in (b), an agent but with respect to the main verb in (c)). However, what Burzio does not say is that the sentences represent cases that are ordered with respect to their position along the configurational hierarchy in Table 3.1: (a) is a core case, (b) is a peripheral case, and (c) is an optional application of a restructuring rule. It appears that the possibility and relevance of a semantic characterization of auxiliary selection is directly proportional to the centrality of verbs along the hierarchy. Thus, core cases like (a) belong to a conceptual class of predicates whose common denominator is 'existence or change of state'. It will be argued later in this chapter that such a class has internal distinctions (semantically compatible with the common denominator) that are relevant from the point of view of language change and language acquisition. Peripheral cases like (b) and (c), on the other hand, are not easily characterizable semantically. This therefore brings us back to the fundamental distinction, within the domain of auxiliary selection, between phenomena that lend themselves to a characterization in semantic terms and phenomena that can only be accounted for syntactically. The success of a syntactic theory is often measured by the extent to which it encompasses all the relevant phenomena without resorting to semantics. By this metric of evaluation, GB theory successfully achieves this result, but at a high level of abstraction from details. The price to pay is the loss of explanatory power with

respect to a range of subtle differences in the behaviour of individual verbs or verb classes.

For example, it has been remarked (see Centineo 1986) that Burzio's theory deals inadequately with verbs like *correre* ("run") that can take either auxiliary. As it stands, the theory is unable to establish a principled relationship between an unergative verb with a thematic subject position and no object, and an unaccusative verb with no thematic subject and a filled direct object position. Verbs like *correre* would therefore have to be entered in the lexicon twice, once as an unergative verb, and once as an unaccusative verb. As the next section will show, some semantic theories are better able to account for this problem.

A more fundamental drawback of GB theory, and of any purely syntactic account of auxiliary selection, is that it does not really explain why unaccusatives select ESSERE and unergatives select AVERE (see Grimshaw 1987 on this issue). The rule for auxiliary selection, in both Relational Grammar and GB, is based on a relationship between object and subject positions, which puts unaccusatives, reflexives and passives into the same category. What remains unexplained is why this particular relationship between subject position and object position should determine auxiliary selection, whereas other relationships between positions are irrelevant; and why it is ESSERE that is selected when the relationship obtains, and not AVERE. As we saw in section 3.3, this is also a problem for the construct of Multiattachment in Relational Grammar.

3.8 Lexical-semantic aspects of unaccusativity

It is fair to say that purely syntactic approaches fail to distinguish the whole class of unergative verbs from the whole class of unaccusative verbs cross-linguistically: in many languages (other than Italian), unaccusative verbs are cross-classified by syntactic diagnostics. For example, ESSERE-assignment distinguishes unergative and unaccusative verbs in Italian, but ETRE does not have this function in French. Stating the differences between French and Italian in syntactic terms does not lead us very far in understanding the nature of such differences. The syntactic hierarchy proposed in Government-Binding terms identifies a continuum of configurations, some of which are more central to ESSERE/ETRE selection than others. Yet the question of precisely what factors are responsible for the variation among French non-reflexive unaccusatives is not

addressed. Burzio argues that 'lexical factors' are responsible for the inconsistent behaviour of French unaccusative verbs with respect to auxiliary selection, but does not explore the reasons underlying such inconsistency. Similarly, Pearce (1990: 106-7) argues that the possibility of ETRE with unaccusative verbs is 'lexically conditioned'. Underlying both definitions seems to be the assumption that 'lexical' means 'unsystematic' and 'not amenable to systematization'.

It is clear that the important generalizations established by purely syntactic approaches appear to be constrained in the vast majority of cases by semantic factors. This is not a serious challenge to the Unaccusative Hypothesis (see Rappaport 1989 on this point): it suggests that syntactic configurations are a necessary but not sufficient condition of unaccusativity/unergativity, and requires consideration of the mutual influence of syntactic and semantic factors.
As Guerssel et al. (1985) put it, native speakers of a language have knowledge of 'syntactically relevant semantically coherent verb classes'. The relevant question from the point of view of language acquisition is how native speakers acquire knowledge of these verb classes. Pertinent questions from a second language acquisition perspective are also whether, and to what extent, non-native speakers come to acquire such knowledge. In order to address these issues, one has to look at the complex interplay of underlying semantic factors and overt syntactic realizations.

Various theoretical accounts have dealt with the question of unaccusativity and auxiliary selection from a semantic point of view. These theories can be grouped according to the following distinctions:

(a) some theories concentrate on the thematic characterization of the arguments of verbs selecting either one or the other auxiliary; others focus on the aspectual dimension of the event structure associated with these verbs. A third group of theories look at the argument structure associated with particular verbs as an interface between the lexical-semantic feature of verbs and their syntactic properties.

(b) more generally, some theories stress the adequacy of a purely semantic approach to auxiliary selection, and downplay the relevance of syntactic approaches. Other theories set out to examine the interplay between syntactic and semantic aspects.

The overview in the following section aims to demonstrate that purely semantic accounts of auxiliary selection are as inadequate as purely syntactic accounts, although for different reasons. It will be argued that the best account is one that looks at how the semantics of a verb determines the syntactic character of its arguments, which in turn affects the syntactic behaviour related to unaccusativity and auxiliary selection. It will be shown that (a) a finer-grained analysis of the semantics of verbs reveals important differences within the classes of unaccusative and unergative verbs, and that (b) these differences affect both the way verb categories are perceived by native speakers of Italian and the way they are acquired by speakers of other languages.

3.8.1 The Universal Alignment Hypothesis

It has been mentioned before that the early formulation of the Unaccusative Hypothesis within the framework of Relational Grammar addresses the question of how and to what extent initial grammatical relations correlate with semantic roles. Perlmutter and Postal (1983: 97) make a strong claim about the link between initial relations and semantic roles in what they call the 'Universal Alignment Hypothesis', stated here in (101) :

(101) There exist principles of universal grammar on the basis of which, given the semantic representation of a clause, one can predict which initial grammatical relation each nominal bears.

If this hypothesis is correct, then one could infer the assignment of initial intransitive clauses to the unergative or to the unaccusative sub-type from their meaning. This entails that clauses that have the same meaning must have the same initial strata (i.e. either unergative or unaccusative) in all languages.

A necessary first step in the direction of testing the Universal Alignment Hypothesis is the identification of the semantic factors that, cross-linguistically,correlate with initial unergativity vs initial unaccusativity. Perlmutter and Postal (1983: 98) arrive at a preliminary classification that singles out types of predicates determining unergative and unaccusative initial strata. According to this classification, unergative strata tend to correspond to the traditional notions of 'activity', which includes (at least) two subcategories:

(a) predicates describing willed or volitional acts, including 'manner of speaking' verbs (i.e. *work, play, speak, smile, etc.*) and 'predicates describing sounds made by animals' (i.e. *bark, neigh, quack,* etc)
(b) predicates describing certain involuntary bodily process (i.e. *cough, sneeze, cry,* etc.)

The class of predicates determining initial unaccusativity includes:
(a) predicates whose initial nuclear term is semantically a patient (i.e. *burn, fall, slide, drip,* etc);
(b) predicates of existing and happening (i.e.*exist, happen, occur, arise,* etc);
(c) predicates denoting involuntary emission of stimuli that impinge on the senses (i.e. *shine, glow, sparkle*, etc);
(d) aspectual predicates (i.e. *begin, stop, start,* etc.);
(e) durative predicates (i.e. *last, remain, stay, survive,* etc.).

This classification is, by admission of its authors, tentative and incomplete: it does not rely on consistent criteria (it makes unsystematic reference to both thematic and aspectual dimensions); it does not attempt to find a common denominator to the various verb classes; moreover, it is not related to an explicit theory of the mapping between semantic and syntactic representation. But the most serious problem of all is that such imprecise formulation, coupled with the ambitious aim of the Universal Alignment Hypothesis, makes the hypothesis extremely vulnerable to counterevidence: indeed, any piece of evidence that shows that clauses with the same meaning do NOT have the same initial representations in different languages counts as a a counter-example. The study by Rosen (1984), perhaps the most devastating criticism of the hypothesis, abounds with examples where a particular semantic role maps onto an initial 1 in one language and an initial 2 in another: the predicate 'sweat', for example, has an initial 1 in Italian but an initial 2 in Choctaw. It follows that the strong form of the Universal Alignment Hypothesis is untenable, i.e. that initial representations cannot be determined on the basis of meaning alone.

It would be difficult to deny, however, that there are obvious cross-linguistic tendencies for certain meanings to be associated with initial unergativity or initial unaccusativity: these tendencies require an explanation. Keenan (1988) presents an overview of cross-linguistically valid semantic properties of *absolutives* (i.e. subjects of intransitive verbs and direct objects of transitive verbs). Such properties

are characterized in terms of bondedness to the verb, thematic role, and control phenomena although, as Keenan himself emphasizes, the distinction is somewhat artificial, given that the three categories overlap to a considerable extent.

'Bondedness to the verb' refers to the tendency of referents of absolutives to appear to come into existence through the activity expressed by the predicate. In contrast, the referent of transitive subjects is understood to exist independently of such activity. So in (102) below, the existence of the puddle is not independent of the act of forming:

(102) A puddle formed on the floor

The close link between the subject of absolutives and the predicate is also evident in the selectional restrictions on the kind of absolutive arguments that can undergo the activity specified by the verb. For example, the sort of things that can spill (intransitive) are the sort of things that someone can spill: thus, they are limited to liquids or similar objects (beans, etc.).

The thematic properties identified by Keenan correspond to the common thematic roles of Agent, Patient, Theme, etc. Absolutives are usually patients, in the sense that their existence state is understood to be affected by the action expressed by the predicate. Moreover, absolutive arguments of verbs of motion have their path of movement specified by Goal and Source locatives, in the sense indicated by Gruber (1976) and Jackendoff (1972). Themes can be either 'objects which move' in the concrete sense, or metaphorical extensions departing from the concrete sense.

Control phenomena refer to a class of expressions, such as adjectives and infinitival phrases, which occur with predicates and are understood to predicate something of the argument. Thus, in (103a), the adjective expresses a property of the referent of the subject argument; in (103b), the predicate denotes a change of state in the argument, which is understood as having acquired the property expressed by the predicate:

(103) a. John looks happy
b. The milk turned sour

The same control phenomena are exhibited by transitive verbs with respect to their direct object.

Languages have productive means of deriving one-place predicates from two-place predicates which "respect absolutives": this means that the absolutive argument of the derived predicate (a subject) has the same properties as the absolutive argument of the predicate it is derived from. Clearly, this is consistent with the assumption underlying syntactic theories of auxiliary selection about the derivation of unaccusatives from transitive verbs: subjects of unaccusatives behave like objects of transitive verbs.

Cross-linguistic data therefore confirm the systematicity of the semantic distinctions between unergativity and unaccusativity, albeit in the form of statistical tendencies and not absolute universals.

Even if one remains skeptical about the universality of the syntax-semantics correlates of unaccusativity, it would be difficult to deny that Italian is a language in which, in the overwhelming majority of cases, grammatical relations ARE predictable on the basis of the semantics of the predicate. This is recognized by Rosen in what she calls the "Little Alignment Hypothesis', which refers only to language-internal evidence without making any universal claims:

(104) The Little Alignment Hypothesis (Rosen 1984: 53)
For any one predicate in any one language, there is a fixed mapping which aligns each semantic role with an initial GR. The alignment remains invariant for all clauses with that predicate.

In spite of a few counterexamples, the Little Alignment Hypothesis in Italian is - if not invariably true - *nearly* true: the selection of AVERE or ESSERE, as visible manifestations of initial unergativity or unaccusativity, correlates with semantic roles according to the predictions.

To sum up, crosslinguistic evidence shows that the Universal Alignment Hypothesis does not have universal validity, but there are significant tendencies across languages that provide support for the predicted correlation between semantic roles and syntactic relations. Language vary as to the degree of internal consistency between the two: Italian is an example of maximally consistent language.

Faced with the lack of absolute reliability of the semantic bases of unaccusativity and auxiliary selection, one might be tempted to reject any semantic characterization of these grammatical phenomena and focus exclusively on syntactic characterizations. This is the path chosen in the Government and Binding approach and in the latest Relational Grammar approach. Alternatively, one could take a more realistic stance and give up expecting a perfect and universal correlation between syntax and semantics within this domain. One could therefore expect to find an indirect relationship between the semantic features underlying classes of arguments, or predicates, and the syntactic behaviour of verbs as unergatives or unaccusatives: such indirect relationships are nevertheless likely to be regulated by principles, which could be at least in part language-specific. As noted by Grimshaw (1987), two steps of argumentation are required: the first is discovering which semantic verb classes act as syntactic unergatives or unaccusatives with respect to the set of well-known grammatical properties, the second involves discovering the mappings of semantic representations onto syntactic configurations.

We will now devote some attention to the interface between the syntax and the semantic of auxiliary selection.

3.8.2 Historical evolution of auxiliaries in Romance

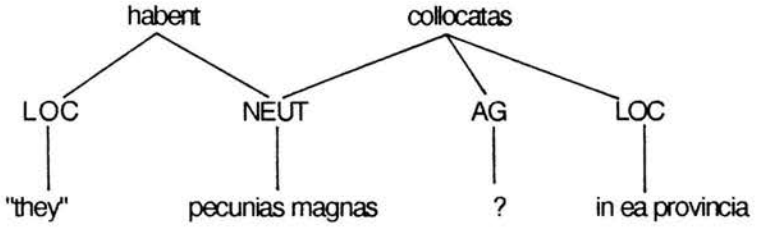
One of the keys to an understanding of the complex interaction of semantic and syntactic factors involved in auxiliary choice is an appraisal of the historical evolution of auxiliaries. Like other instances of language change, the diachronic development of auxiliaries combines gradual semantic changes and abrupt syntactic reanalysis.

Vincent (1982) and Tuttle (1986) provide a semantic characterization of the progressive replacement of inflected forms by periphrastic forms in the evolution of auxiliaries HABERE and ESSE from Latin to Romance. Like most theoretical discussions in the literature, their accounts are concerned with the evolution of the HABERE periphrasis: the ESSE periphrasis has generally been treated as the 'poor relation'.

Using a double system of verbal classification based on both grammatical relations (subject and object) and case relations (Agent, Locative, Experiencer, Neutral), Vincent retraces the process of grammaticization through which the Latin constructions with HABERE and ESSE developed from their original meaning and

uses into the perfective meaning. Originally, HABERE was a two-place verb taking Locative as subject and Neutral as object; ESSE was a one-place verb taking Neutral as its subject; the participle was an adjectival form co-occurring with Neutral. The sentence in (105) exemplifies the HABERE construction than eventually gave rise to the Romance periphrastic perfects:

(105) In ea provincia pecunias magnas collocatas habent (Cicero)
'In that province capital great invested (they) have'



The verb *collocare* takes three arguments: Agent (subject), Neutral (object) and Locative (prepositional phrase). Since the participle modifies a Neutral, *collocatas* in (a) must receive a passive interpretation. Notice that the Locative subject of HABERE is expressed by the inflection, but the Agent subject of *collocare* is not expressed: this allows two possible interpretations of (105), depending on whether the investors of the money and the current possessors are the same people or not.

Consider now a further example:

(106) (Equitatum) quem ex omni provincia coactum habebat (Caesar)
'(cavalry) that from each province gathered (he) had'

This example makes it clear that the Locative of *habere* and the Agent of the participle are not necessarily the same (although they could plausibly be in the above example). To quote Vincent, "...in most instances the circumstances will dictate the identification of the Locative of *habere* with the Agent of the participial verb, and it is then but a short step for this habitual identification to become a grammatically required one." (Vincent 1982: 84). The grammaticization of *habere* involved the retention of the grammatical roles of subject and object, with the concomitant loss of the semantic role of Locative.

The change was initially favoured in the case of transitive verbs where the Agent and the Locative can only be identical, such as the *verba sentiendi* (like *cognosco*, 'I know', *comperio* 'I discover'), denoting 'sensory-intellective' processes (Benveniste 1968: 87) with an Experiencer subject . It then spread to transitive verbs with non-human or inanimate subjects (which could not easily satisfy the Locative function that would have been required by the original paradigm of HABERE of possession), to absolute (objectless) uses of transitive verbs, and eventually to one-place intransitives. The common denominator of all these verbal categories is the agentive, affecting role of the subject.

What the HABERE periphrasis could not accomodate, however, was the class of intransitive verbs whose subjects expressed the role of patient, and/or were affected, rather than affecting. These therefore developed an alternative periphrasis with an auxiliary that required itself a patient subject, namely ESSE. Initially, the ESSE periphrasis came in fact to be associated with the passive construction, which typically involved Neutrals as subjects, and which was already common in Classical Latin. Another category of verbs fitting the Neutral-subject definition was that of deponents, which tended to have a movement or change-of-state element of meaning and generally denoted personal subject involvement. This pattern favoured the extension of periphrastic expressions with *esse* + participial form to other non-deponent verbs of similar meaning (*venio -->ventus sum* as the preferred periphrastic replacement for *veni*), and to 'medio-passives' (i.e. *cingor* 'I gird myself', *memoror* 'I remember (to myself), *induor* 'dress myself').

The development of periphrastic forms in Late Latin, however, was not just a gradual process of lexical diffusion: it has to be considered against the background of other important syntactic changes in the Latin system.
As Ramat (1982) puts it,

> "On the one hand, there was a general tendency to adopt analytical forms in both noun and verb morphology and, on the other hand, a corresponding crisis in the system of inflections with less attention being paid to agreement (e.g. adjective and noun agreement) - which is typical for languages with a high incidence of inflections. Finally, a fixed SVO word order was established. Yet, at the same time, the morphological means which led to the reanalysis were already present in Latin, so that we may speak of a kind of stability of forms within morphological discontinuity."
> (Ramat 1982: 141-142)

The diachronic change that led from 'HABERE + (Object + participle) to '(HABERE + participle) + Object' spread gradually through semantically-defined verb classes. At some point, however, the grammar must have been reanalysed in order for the old construction to be eliminated. In order to understand how the Latin construction turned into the Romance one, it is therefore essential to examine the interplay of syntactic and semantic changes. According to Salvi (1987),

> "The change began at the semantic level: the Latin structure changed its original meaning into the meaning of the Romance periphrasis and the new meaning of the construction only later imposed a syntactic restructuring, which led to the structures of the modern Romance languages."
> (Salvi 1987: 229)

The two main semantic factors that started off the process of syntactic reanalysis were (a) the semantic emptying of HABERE, and (b) the frequent coincidence between the subject of HABERE and the subject of the participle (a coincidence that, as we saw earlier, was normal with verbs expressing intellectual activity).
The first syntactic consequence was an increase in the range of participles found in the construction: originally, only participles of verbs with a resultative meaning, which were able to take an adjectival value, could be found in it. Then, the participle lost its adjectival character and assumed a verbal one. Consider the contrast in (107):

(107) a. multa bona bene parta habemus
'many goods well procured have we'

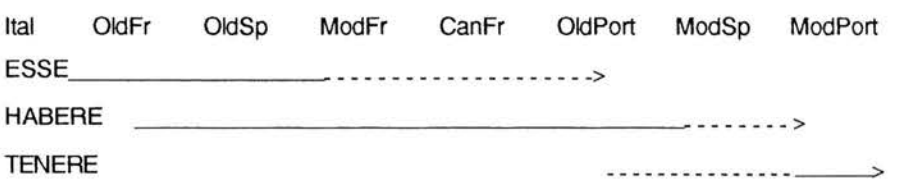
b. haec omnia probatum habemus
'these all tried have we'

In (107a), the past participle still has adjectival character and represents an optional attribute of the noun, with which it agrees, while HABERE has full verbal meaning. There is no necessary connection betwen the subject of HABERE and the logical subject of the past participle. In (107b), the past participle has verbal character and is fully integrated with the auxiliary. One of the surface indicators of reanalysis is the loss of agreement of the past participle with the noun object. HABERE has been semantically 'bleached' and is now simply an auxiliary. Moreover, the subject of HABERE can only be the same as the subject of the past participle.[30]

As Salvi (1987) points out, in the original construction the axis of the semantic interpretation was HABERE with its complements; with the semantic emptying of HABERE the axis is transferred to the participle. Thus, the original construction which signified the possession of the results of an action finally came to indicate the past action itself. But the new semantic interpretation required the fundamental syntactic changes we have just described, which must have occured through a succession of discrete stages.

We have examined the processes that led to the establishment of periphrastic constructions with auxiliaries in some detail, because an understanding of the syntax/semantics interface in these processes may help to shed light on another major diachronic fact: the pan-Romance drift away from the derivations of *esse* in favour of the forms derived from *habere*. Romance languages can be arranged along a continuum defined by the degree of innovation in this respect:

TABLE 3.3 The evolution of auxiliaries in Romance (from Vincent 1982)



Italian is the most conservative as it retains a fully productive contrast between ESSERE and AVERE, whereas French shows a more advanced evolution towards the extension of the scope of AVOIR at the expense of ETRE (and in other varieties of French , like Canadian French, such extension is even more marked); in Spanish and Portuguese the process was completed by the end of the 16th century, and reflexes of TENERE have (to a different extent in the two languages) replaced HABERE-reflexes.

Vincent (1982: 96)) suggests that "...we find here evidence of a gradual syntactic change of a kind that some theoreticians - notably Lightfoot (1979) - have sought to deny." However, it seems to us that what we have evidence for is the gradual spreading of a *lexical-semantic* change. In Chapter 1 we mentioned Lightfoot's distinction between gradual change in the environment and abrupt syntactic

reanalysis in the individual grammars: changes in the linguistic environment or in the linguistic community may be continuous and span over a long time, until syntactic restructuring becomes necessary. We shall see in section 3.10.4 that the replacement of ESSE-reflexes by HABERE-reflexes in Romance can be regarded as a gradual process with lexical-semantic connotations: but the details of the syntactic reanalysis that has been leading to a single-auxiliary system in Romance are still largely unexplained.

3.8.3 Thematic relations in auxiliary selection

3.8.3.1 Thematic relations

As we have seen from the brief account of the historical process of change of auxiliaries, explanations are often couched in terms of thematic roles.
Vincent (1982, 1987), for example, maintains that what is crucial to an understanding of the process of auxiliary selection is an appreciation of the semantic case relation between the subject NP and the verb: intransitive verbs with 'neutral' or 'patient' ("inert") subject NPs take ESSERE whereas those with 'agent' or 'experiencer' subjects take AVERE. Thus, unaccusative verbs of motion, in his view, require ESSERE because the subject is neutrally involved in the state defined by them - in traditional terms a patient.

The problem with thematic roles is that they tend to be overgeneral.[31] Consider, for instance, the term 'neutral': it covers a numer of case relations which had received a different label and a somewhat different content in the literature. In particular, it includes the notion of 'theme' - the NP whose location or change of location is predicated.
The category of verbs with a subject NP in the 'neutral' case is a large one: it minimally includes all verbs of 'change of state' and 'existence of state'. Note, furthermore, that Vincent's subcategory 'verbs of motion' is a misnomer. Some motion verbs (*camminare, nuotare*) take only AVERE while others (*correre, volare*) take AVERE or ESSERE depending on meaning. Those verbs of motion that take ESSERE are, more strictly speaking, verbs of 'change of location', verbs which express an event marking the inception of a NEUTRAL + LOCATIVE relation. As Centineo (1986) points out, it is difficult to maintain that the subject in a sentence like (108) is 'semantically inert", since it appears to be the volitional initiator of the action and

therefore an agent (although it is also a theme, in the sense that it undergoes motion):

(108) Paola è corsa a casa
'Paola ran home'

Vincent's framework cannot distinguish between agent and theme roles, in those cases in which a single NP can express both. There are no principled independent criteria for deciding which roles are assigned to a verb and therefore for explaining auxiliary selection, which is ultimately dependent on such roles.

Vincent's subcategories also fail to include 'impersonal' verbs (i.e. NEUTRAL subject verbs; *sembrare, piacere*, etc.) denoting possessive, emotive, cognitive, perceptual relations subsumed by the case frame [___NEUT, EXPER], which in the context of the present study are seen as falling into the category of verbs which denote the existence of a state.

We shall return to the issue of thematic roles in section 3.9 and show that thematic roles acquire more significance if they are considered in the context of (a) their representation in conceptual structure, and (b) the modes of their projection onto syntactically relevant variables in argument structure.

3.8.3.2 Aspectual relations and auxiliary selection

Other theories of auxiliary selection stress the relationship between auxiliary choice and the aspectual features of verbs. One of the first accounts that was proposed in this framework is Parisi (1976).

Essentially, for Parisi, who adopts a generative semantic type analysis, the factor that determines auxiliary choice and participle agreement is the semantic character of the 'situation' denoted by the verb in question - the key notion for Parisi being whether the verb denotes a STATE or CHANGE OF STATE as opposed to an ACTIVITY or PROCESS. As to past participle agreement, he identifies two main classes of verbs:

Category A: verbs that imply a (change of) state in one of their arguments:

(109) Maria è uscita
'Mary has gone out'

(110) Maria ha lavato le finestre
'Maria has washed the windows'

Category B: verbs that do not imply a change of state in any of their arguments:

(111) Maria ha dormito
'Maria has slept'

Within category A, the past participle will agree with the subject only if the verb's state is predicated of the subject. If this argument is selected as direct object, the standard usage prefers the non-agreement of the past participle. There is, however, a sub-standard usage which allows for the participle to agree with the direct object:

(112) Maria ha lavate le finestre
'Maria has washed the windows'

Within category B, which does not imply a change of state, the past participle remains unaltered.

Auxiliary choice, in Parisi's approach, is governed by the following aspectual constraints:

(a) if the predicate, either directly (as in adjectives), or indirectly (as in past participles), expresses a change of state in the argument selected as subject, the auxiliary verb is ESSERE;

(b) if the predicate does not express a change of state in the argument selected as subject, either because it does not express a state at all (as in *dormire*, 'sleep') or because it expresses a state in the argument selected as direct object (as in all transitive verbs), the auxiliary verb is AVERE.

In Parisi's view, this accounts for the ambiguity of verbs such as *correre* ('run') which takes ESSERE when it denotes a change of state in the subject (i.e. a change of location) but takes AVERE when it denotes an activity:

(113) Paola è corsa a casa di Piero
'Paola ran to Piero's house'

(114) Paola ha corso velocemente
'Paola ran quickly'

Similarly, verbs like *aumentare* ('increase') may be either intransitive or transitive, depending on whether they denote a change of state in the subject argument or the causation of a change of state in the object argument:

(104) I prezzi sono aumentati [change of state in SU]
'Prices have increased'

(105) I negozi hanno aumentato i prezzi [change of state in DO]
'The shops have increased the prices'

It can be seen that Parisi's criteria separate well-defined categories (including verbs that only allow one interpretation, and therefore enjoy unambiguous membership) from more indeterminate categories (which typically include verbs corresponding to more than one interpretation). Given that the Unaccusative Hypothesis was not available to him, however, Parisi could not exploit the relevance of the systematic relation between objects of transitive verbs and subjects of intransitive verbs.

It should be apparent that all change-of-state verbs satisfy Burzio's syntactic definition of unaccusativity. However, the notion of "change of state" which underlies Parisi's classification appears to be too broadly defined. The class of intransitive verbs which may only take ESSERE as auxiliary does not only include verbs denoting a change of location or condition in the subject argument (such as *andare, venire, diventare*, etc): it also includes verbs denoting the continuation of a state, as in (117a,b),, and verbs denoting the mere existence of a state, as in (118a,b):

(117) a. I miei nonni sono sopravvissuti alla guerra
'My grandparents survived the war'
b. Maria è rimasta a casa
'Maria stayed at home'

(118) a. Questa pipa è appartenuta a mio padre
'This pipe belonged to my father'
b. Quella storia mi è sembrata incredibile
'That story seemed incredible to me'

The other interesting issue that arises out of Parisi's account is that of the optional agreement of the participle with the direct object of transitive verbs. Whether such agreement is possible or not depends on the resultative character of a verb, and by the consequent ability of the participle to denote a new state. As we saw earlier, past participle agreement with the object of resultative verbs was the fundamental feature of the original Latin construction involving HABERE in its full verbal meaning.
It seems, as Parisi suggests, that past participle agreement is more acceptable with transitive verbs implying the causation of a change of state in the object than with transitive verbs not implying such a change of state: in terms of Vendler's (1967) classification, it is possible with accomplishments but not with achievements. Yet there is evidence from first language acquisition indicating that Italian children initially make achievement (and not accomplishment) verbs agree with the noun:

(119) Presa Checco campana
taken Checco bell
'Checco took the bell'
(example from Antinucci and Miller1976)

This example suggests that aspect - rather than tense - is of primary importance from the point of view of children's cognitive development (see Ramat 1982 for comments on this issue; for a syntactic explanation, see the recent GB analysis by Borer and Wexler 1992).[32]
As will be seen in Chapter 5, the acceptability of past participle agreement with the direct object of resultative and non-resultative verbs was one of the questions addressed by a pilot study that tested the linguistic intuitions of native and non-native speakers of Italian.

3.8.3.3 Role and Reference grammar

The Role and Reference grammar (RRG) framework (Foley and Van Valin 1984; Centineo 1986, Van Valin 1990) attempts to reconcile the two levels of thematic and aspectual analysis.

RRG differ from other theories of auxiliary selection because it posits only a single level of syntactic representation. The essential components of this approach are a system of verb classification and predicate semantics (Vendler 1967; Dowty 1979) and a theory of thematic roles. Foley and Van Valin's starting point is the system of verb classification and semantic decomposition developed by Dowty (1979) on the basis of Vendler's classification of English predicates into *activities, accomplishments, achievements, and states*. This system is reported in Table 3.4 below:

TABLE 3.4. Verb classes and logical structures

Verb class	Logical structure
State	predicate' (x)
Achievement	BECOME predicate ' (x)
Activity	DO (x, [predicate' (x)])
Accomplishment	[DO (x, [predicate' (x)])] CAUSE [BECOME predicate'(y)]

The logical structures in Table 3.4 are the basis for a theory of semantic roles which includes two tiers: one corresponding to the thematic relations of other theories, and another which has no exact counterpart in other theories.

The derivation of thematic relations from argument positions (the first tier of the theory) is summarized in Table 3.5:

TABLE 3.5: Logical structures and thematic relations (adapted from van Valin 1990)

I.	State verbs		
	A. Locative	be-at' (x,y)	x=locative y=theme
	B. Non-locative		
	1. State or condition	predicate' (x)	x=patient
	2. Perception	see' (x,y)	x=experiencer y=theme
	3. Cognition	believe' (x,y)	x=experiencer y=theme
	4. Possession	have' (x,y)	x=locative y=theme
	5. Attrib/identificational	be' (x,y)	x=locative y=theme
II.	Activity verbs		
	1. Controlled	DO(x, [predicate' (x)])	x=agent (y=locative)
	2. Uncontrolled	predicate' (y)	y=effector (y=locative)

The various thematic relations are subsumed under the generalized semantic 'macroroles' of *actor* and *undergoer*, which constitute the second tier of semantic relations. The actor is "the argument of a predicate which expresses the participant which performs, effects, instigates, or controls the situation denoted by the predicate" (Foley and Van Valin 1984: 29): the notion includes the semantic relations of agent, effector and experiencer. The undergoer is "the argument which expresses the participant which does not perform, initiate or control any situation but rather is affected by it": the notion subsumes the semantic relations of patient and theme (and locative with some predicates).

The interpretation of a relation as actor or undergoer is regulated by the two accessibility hierarchies in (120a), (120b), which can be combined into the single hierarchy reported in Table 3.6:

(120)

a.	Actor	agent > effector > experiencer
b.	Undergoer	patient > theme > locative

TABLE 3.6 Actor/Undergoer Accessibility hierarchy

Actor					**Undergoer**
Agent	Effector	Experiencer	Locative	Theme	Patient
---------->					
		<----------			

The direction of the arrows indicates the increasing markedness of the realization of thematic relations as macroroles.
In a transitive sentence there can only be one actor and one undergoer. The situation is more complicated for intransitive sentences,since the single argument of the verb will be either an actor or an undergoer depending on the semantic class of the predicate, i.e. whether the latter is an activity or an achievement.

Role and Reference grammar employs the concept of "pivot of a syntactic construction" instead of subject and direct object. Pivots are not necessarily syntactic, although in many languages most of the major syntactic constructions have the same pivot. In Italian, as in other languages which have syntactic pivots, the pivot corresponds to the traditional notion of syntactic subject. In the passive construction, the actor is the unmarked pivot choice and the undergoer is the marked choice.

The application of this rather complex semantic apparatus to the analysis of auxiliary assignment in Italian yields two explanations: one relies on the inherent aspectual properties of verbs (see van Valin 1990 for an exhaustive summary); the other one emphasizes the Actor/Undergoer accessibility hierarchy (se Centineo 1986).

From an aspectual point of view, the generalization is that in Italian all AVERE verbs are activity verbs and all ESSERE verbs are state, achievement, or accomplishment verbs. Given the logical structures in Table 3.4, one can say that ESSERE verbs have a 'state' element in their semantics, while AVERE verbs do not.
Among ESSERE verbs, the state class includes predicates describing conditions of being (*esistere*), locative predicates (*stare, rimanere*), predicates of possession or perception (*piacere, appartenere, bastare*); the achievement class includes inchoatives *(migliorarinvecchiare*) and verbs of happening (*succedere, accadere*);

Verbs of motion belong either to the achievement class (e.g.*arrivare*)or to the accomplishment class (e.g.*andare*). Other verbs of motion (e.g. *correre*) exhibit an alternation between activity and accomplishment semantics: they are accomplishments whenever the source and/or location is specified, and activities when they occur without a goal adverbial (*'correre a casa* 'run home') is classified as an accomplishment, whereas '*correre velocemente* 'run fast') is an activity). Paired unaccusative verbs with a transitive alternant (e.g. *aumentare, migliorare*) are classified as intransitive achievements, and their alternants as transitive accomplishments. "Weather" predicates (which may occur with either auxiliary in apparent free variation) may express an activity or an accomplishment, as in (121):

(121)	a.	Ha/è piovuto incessantemente	activity, ESSERE/AVERE
	b.	Mi è piovuto sulla testa	accomplishment, ESSERE
	c.	*Mi ha piovuto sulla testa	accomplishment, AVERE

Auxiliary selection and NE-cliticization, according to this theory, have the common requirement that the verb have a state predicate in its logical structure. The rules for the two phenomena can be formulated as follows (see van Valin 1990: 233 for details):

(a) AUXILIARY SELECTION IN ITALIAN: select ESSERE if the logical structure of the verb contains a state predicate;
(b) NE-CLITICIZATION: NE realizes the lowest ranking argument on the Actor-Undergoer hierarchy in the state predicate in the logical structure of the predicate.[33]

The generalization that AVERE verbs are activity verbs and intransitive ESSERE verbs are either state, or achievement, or accomplishment verbs can be further specified by analysing the realization of thematic roles as macroroles. In particular, it appears that

(a) activity verbs have an actor (an agent, an effector, or a locative) as pivot;
(b) state, achievement and accomplishment verbs have an undergoer (a patient or a theme) as pivot.

It follows that within this framework all verbs satisfying Burzio's criteria for unaccusativity have an undergoer argument. Since they all pass the NE-cliticization

test, the latter is regarded as a test for undergoer pivot. Assuming an undergoer pivot for ESSERE-verbs, however, does not solve the problem of verbs whose pivot exhibits properties of both actor and undergoer (Vincent's neutral subject). In sentences like (108) - repeated below as (122) - the same argument expresses both the volitional initiator of the action and the participant which undergoes a change of location.

(122) Paola è corsa a casa
'Paola ran home'

The pivot of verbs like *correre* (and *saltare, volare, strisciare*, etc.) is what may be called an affected actor. Affected arguments, i.e. theme and patients, are highly marked choices for the role of actor. This accounts not only for unaccusative ESSERE verbs but also for passive constructions and benefactive reflexives, which clearly have an affected actor as pivot.

The global generalization that can be drawn is that ESSERE is selected by those verbs that have a marked choice as pivot, i.e. an undergoer, or by those with a non-prototypical pivot, i.e. an affected actor. AVERE, on the other hand, is selected whenever the pivot is unmarked, i.e. an actor, and prototypical, i.e. it is affecting but not affected.

The effect of markedness on pivot choice and auxiliary assignment is shown by Table 3.7:

TABLE 3.7. Continuum of markedness for pivot choice and auxiliary assignment

	ACTOR	Transitive	Accomplishments	least MARKED	A
	(-affected)	"	Activities		V
P		"	Achievement		E
		"	States		R
I		Intransitive	Activities		E
V	(+affected)	Transitive	Reflexives		E
		Intransitive	Accomplishments		S
O		"	Achievements		S
		"	States		E
T					R
	UNDERGOER	Passives		most MARKED	E

The semantic analysis of Italian auxiliary within Role and Reference grammar therefore confirms the markedness of ESSERE.[34] The analysis combines various

elements of the other approaches, occasionally compensating for various deficiencies. It does, however, generate a proliferation of distinctions that, given the inherent fuzziness of some of the semantic categories (such as achievements and accomplishments with respect to ESSERE verbs of motion), will either spawn more subcategories or succumb to vagueness. Van Valin's conclusion that the lexical semantic account is superior to any other certainly appears overstated. What is needed is a further effort towards the integration of syntactic and semantic approaches. Research on argument structure, which is the topic of next section, in our view represents the best attempt to integrate the two sides.

3.9 Argument structure and auxiliary selection

3.9.1 General remarks

Ever since Chomsky's "Remarks on Nominalization" (1970), the lexicon, and lexical representation, have assumed an increasingly central role in syntactic description. Research on argument structure has expanded dramatically in the last ten years, because of the increasingly important role played in Government-Binding by principles such as the Theta Criterion and the Projection Principle, and also because of the appearance and development of lexicalist theories like Lexical Functional Grammar (Bresnan 1982).

Argument structure represents a complex of information crucial to the syntactic behaviour of a lexical item. It interacts with two other kinds of representation: on the one hand with lexical-semantic structure, which represents lexical meaning, and on the other hand with deep structure (d-structure). Argument structure is the interface between lexical-semantic structure (also called lexical-conceptual structure) and d-structure: it is projected from the former, and is the input to the latter. According to the strongest prediction, the argument structure of a lexical item is predictable from its meaning, and the d-structure in which the item appears is predictable from its argument structure in conjunction with language-specific parametric characteristics (Grimshaw 1991).

The predicate-argument representation of a verb, or its theta-grid (Stowell 1981), minimally takes the form that indicates the number and types of theta-roles that the verb requires:

(123) PUT: <Agent, Theme, Location>

The labels in this representation are usually taken from the vocabulary of lexical semantics. The notation above is both the most neutral one and the least informative: it provides a list of theta-roles, without making any assumptions as to their mutual relationships. Other notations are found in the literature which assume some internal organization of argument structures. In one of the most common formulation, arguments are distinguished in terms of whether they are external or internal (Williams 1981): internal arguments are in the same phrase as the head verb and are realized in the syntax as direct objects and complements, whereas external arguments are outside the phrase containing the head verb, and are normally realized as subjects. Internal arguments can be further divided (Rappaport and Levin 1986) into direct internal arguments, which receive their theta-roles directly from the verb, and indirect internal arguments, which receive a thematic role from a preposition. Thus, one of the conventions for distinguishing the external from internal arguments in argument structure is the following:

(124) PUT: agent <theme, location>

An alternative notation uses arbitrary variables for the identification of arguments, thus avoiding thematic roles labels altogether:

(125) PUT: x <y, Ploc, z>

Theta-role labels are objects of a lively discussion in the generative literature. While the details of such debate would go beyond the scope of the present exposition, suffice it to say that the vagueness of their definitions, as well as the lack of consensus about their number and the criteria for their application to any given argument, have led a number of researchers to conclude that theta-roles have no status in argument structure (see Zubizarreta 1987; Rappaport and Levin 1986; Grimshaw 1991). In recent theories of argument structure, it is common to make a distinction between two levels of lexical representation: a lexical-syntactic level, which properly encodes the predicate-argument structure of a verb, and a lexical-semantic or lexical-conceptual level, which represents the meaning components of a verb.[35] Theta-roles are not represented in the former, but find a place in the latter. Argument structures are therefore defined only on the basis of

their formal characteristics: they are purely syntactic entities, void of any semantic content.

Conceptual structure does not interact directly with the syntax: rather, it is argument structure that mediates between conceptual structure and d-structure, and that ultimately determines the syntactic configuration of a sentence. Theories of argument structure therefore have to account for a double 'mapping problem': on the one hand, the mapping of arguments at lexical-conceptual structure onto variables at argument structure and, on the other hand, the mapping of variables at argument structure onto positions in d-structure: 'linking rules' of various forms have been posited that establish these correspondences.[36]

It is important to stress that conceptual structure is not equivalent to the entire set of cognitively available distinctions: only a subset of these are relevant to linguistic processes and grammatical encoding. With Pinker (1989: 73), it is possible to say that what enters in the determination of the argument structure is a "thematic core": a schematization of a particular type of event or relationship that constitutes the core of the meanings of a class of possible verbs. The lexical-conceptual structure of intransitive verbs, for instance, has two different thematic cores associated with it: the one underlying unergative verbs, where x performs some activity, and the one underlying unaccusative verbs, where x exists or undergoes some change of state or location:

(126) a. Unergative:
x acts

b. Unaccusative:
x is in a location or state or goes to a location or state

The argument of an unergative verb is agentlike, whereas the argument of an unaccusative verb is a theme. Even the concise semantic characterization offered by the 'thematic core' makes it plain that the argument 'theme' subsumes a variety of semantic events and/or processes. Yet, both agents and themes are mapped onto the surface subject position. We have seen that theories differ as to why this happens. Government-Binding posits movement from an underlying direct object position to an empty subject position; Relational Grammar assumes an analogous

promotion of the theme argument from an initial stratum where it leads a 2-arc to a final stratum in which it leads a 1-arc. Crucially, there is a lack of correspondence between a richly-articulated level of lexical-conceptual structure and the level of argument structure, which allows generalizations that are made possible by ignoring conceptual distinctions. Clearly, the systematicity of the grammatical consequences associated with unergativity vs unaccusativity requires a level of representation at which the two classes of verbs are easily differentiated. This level is provided by argument structure, which captures the syntactic differences between unergative and unaccusative verbs through different configurations where unergatives have a single external argument, while unaccusatives have a single direct argument but no external argument.

For the purposes of the present study, three specific proposals within this area will be examined in some detail. The first one is due to Rappaport and Levin (1986), who assume that an articulated representation of the lexical-conceptual level through meaning decomposition is necessary in order to define semantically coherent classes of verbs, whose members project their arguments onto similar predicate-argument structures. The second proposal is Zubizarreta's (1987), who attributes the irregular behaviour of paired unaccusative verbs in French to a 'marked' mapping between the lexical-semantic level and the syntactic level.
The third proposal is due to Grimshaw (1991), who attributes a structured representation to argument structure which is partly a function of a thematic hierarchy, so that some arguments are more prominent than others. Furthermore, this theory explores the interaction of the thematic and the aspectual dimensions in the definition of structured argument structures: this has the advantage of combining features dealt with in isolation by other proposals, and reaches a more comprehensive explanation of the mechanisms underlying unaccusativity and auxiliary selection.

3.9.2 Predicate decomposition

Rappaport and Levin begin by observing that theta-roles reduce to little more than grammatical functions in the absence of clear criteria for associating specific arguments with specific thematic roles. One therefore witnesses a constant tension between the necessity of providing encompassing generalizations for given syntactic phenomena, and the desirability of accounting for differences in the way individual verbs fit those generalizations. This tension is particularly acute when

lexical-semantic representations consist of an unstructured list if theta-roles because, as Rappaport and Levin point out:

> "...theta-roles are inherently relational notions; they label relations of arguments to predicators and, therefore, have no existence independent of predicators. A list of theta-role labels proves to be inadequate as a lexical-semantic representation since it obscures the complex cross-classification of verbs according to shared components of meaning, which is manifested in various syntactic properties related to the expression of arguments."
> (Rappaport and Levin 1988: 17)

Consider, as an example, the so-called 'locative alternation':

(127) a. John sprayed paint on the wall (*locative* variant)
b. John sprayed the wall with paint (*with* variant)

Although the two sentences in (127) describe the same event (i.e. an entity - a Locatum - coming to be at a particular location - a Goal - through the action of an Agent), they have slightly different semantic interpretations: when the Goal argument is realized as a direct object (as in (127b), it is understood to be wholly affected by the action, whereas this 'holistic' interpretation is not conveyed when the Goal is realized as an object of preposition (as in (127a). Thus, the two variants are only near-paraphrases of each other.

An adequate lexical semantic representation of the locative alternation should be able to account for (a) the near-paraphrase relation, (b) the linking of arguments in terms of their theta-roles, and (c) the 'holistic' interpretation of the Goal as a direct argument. These requirements are not met by a lexical-semantic representation that consists of a list of theta-roles. Suppose, in fact, that the two variants of the locative alternation share a single theta-role list:

(128) SPRAY: <Agent, Locatum, Goal>

The existence of two locative alternation verbs would then be due to two distinct argument structures:

(129) a. SPRAY: x <y, P_{loc} z>

b. SPRAY: x <y, P_{with} z>

This representation is, however, problematic. First, it has to be explained how two distinct argument structures can be derived from a single lexical-semantic representation. The answer may reside in an examination of the linking rules that connect theta-role labels in lexical-semantic representations with variables in argument structures. Such linking rules can be conventionally stated as follows:

(130) LINKING RULES

(a) Link the Agent role with the external argument variable
(b) Link the Theme or Patient role with the direct argument variable
(c) Link any remaining theta-role to an indirect argument variable which is associated with the appropriate preposition.[37]

The Locatum argument, which is the argument denoting the entity that undergoes movement, can be regarded as a Theme (in the sense of Gruber 1965 and Jackendoff 1972, 1983), and therefore linked to the direct argument variable in argument structure, while the Goal would be linked to the indirect argument variable asociated with the appropriate locative preposition. This, however, works with the locative variant, but not with the *with* variant. The problem with the latter is that the Goal role has to be linked to the direct argument variable, thus requiring an idiosyncratic linking rule. Furthemore, a special 'rule of interpretation' has to be posited in order to assign the 'holistic' reading to this particular realization of the Goal role (and to exclude this reading when the Goal is realized as an indirect argument variable).

Alternatively, one could assume that the two variants of the locative alternation have two distinct lexical-semantic representations:

(131) a. SPRAY: <Agent, Theme, Goal> (locative variant)
b. SPRAY: <Agent, Theme, Locatum> (*with* variant)

On this assumption, the Goal in the *with* variant is a Theme is the sense of 'affected entity' that undergoes a change of state (Anderson 1977). The 'holistic' interpretation is then derived from the particular theta-role of Theme associated with the NP. The entity originally described as a Locatum, on the other hand, would keep

this role in the lexical-semantic representation, so that a distinction would have to be formalized between entities undergoing a change of state (Theme)) and entities undergoing a change of location (Locatum).

It is clear that this alternative account creates more problems than it solves. First, it introduces a new theta-role that has not been independently attested in the literature. Second, it does not provide any justification for the fact that, since a clear definition of Theme is lacking, either a Goal or a Locatum role can be linked to the direct argument variable in argument structure; one argument undergoes a change of state and another undergoes a change of location but there is no principled way of deciding which type of change is relevant to the application of linking rules. Third, it is incapable of capturing the near-paraphrase relation between the two variants.
The conclusion is that representations that consist of theta-role lists are inadequate: an analysis that relies on a single theta-role list accounts for the near-paraphrase relation but fails with respect to linking rules and affected interpretation. An analysis based on two theta-role lists accounts for the affected interpretation and linking rules but cannot capture the near-paraphrase relation.

Rappaport and Levin argue that what is needed is a more complex and structured lexical-semantic representation. This involves the central idea of *predicate decomposition* (largely inspired by the work of Jackendoff 1972, 1983): meanings of verbs have internal structure, because they are composed of a set of primitive elements that recur in the definition of many verbs, and characterizes distinct verb classes. These elements can be represented by means of lexical-conceptual structures which may have a number of sub-structures. The different participants in the action denoted by the verb are not identified by theta-role labels, but rather by variables occuring in sub-structures. So, for example, the verb *put* would have the following lexical-conceptual structure:

(132) PUT: [x cause [y come to be at z]]

The entity y, which would be a Theme in the traditional approach, now becomes a variable that occurs in appropriate sub-structures:

(133) a. ...[x come to be at LOCATION]
b. ...[x come to be in STATE]

The relevant linking rule would then refer to these sub-structures, as in (134):

(134) When the lexical-conceptual structure of a verb includes one of the substructures in (133), link the variable represented by x in either sub-structure to the direct argument variable in the verb's argument structure.

In the case of *put*, the linking rule in (134) would apply to the sub-structures in (133), which are subsumed by (132).

In the locative alternation, the locative variant denotes a simple change of location and can therefore be represented in the same way as a verb like *put*. The verb in the *with* variant needs a more complex representation, since it has a component of meaning that the locative variant lacks: this is the bringing about of a change of state of the goal argument, which is at the origin of the 'holistic' interpretation. In fact, the meaning of the *with* variant properly includes the meaning of the locative variant: if John sprayed the wall with paint, then he necessarily sprayed paint on the wall, but not vice versa. What is needed therefore is a lexical-conceptual structure that captures this entailment relation, like (135) below:

(135) a. SPRAY: [x cause [y come to be at z]]

b. SPRAY: [[x cause [z to come to be in a STATE]]
BY MEANS OF [x cause [y to come to be at z]]

(135a) indicates that *spray* denotes an event which involves a change of location (the locative variant), whereas (135b) indicates that *spray* denotes an event which involves a change of state brought about by means of a change of location (the *with* variant).

Unlike the single theta-role list in (128), this analysis can capture the fact that the entity denoted by the z variable both undergoes a change of state (the event in the main clause) and serves as a goal of the change of location of the entity denoted by the y variable (the event in the embedded clause). The use of the same set of variables in both the main and the subordinate clauses makes this possible (unlike theta-role labels, which identify individual semantic roles).

As to linking rules, it may be assumed that, as a general convention, the basic class membership of a verb and the linking of variables are determined by the main clause of the decomposition. Thus, the verb in the *with* variants is primarily a verb of change

of state, and the z variable in (135b), corresponding to the argument denoting the entity that undergoes a change of state, will be linked to the direct argument variable.

This approach emphasizes the fact that single, all-encompassing theta-roles serve a limited function. More richly articulated lexical-conceptual structures allow the expression of the fact that quite different meanings can be associated with each theta-role, and that the identification of such meanings may be necessary in order to account for the syntactic realization of arguments.

However, Rappaport and Levin stress that predicate decomposition does not obviate the need to postulate an additional lexical-syntactic level of representation such as argument structure, because certain generalizations can only be economically stated at that level. The distinction between unaccusative and unergative verbs is a case in point, since many of the phenomena that systematically distinguish the two verb classes can be captured in terms of argument structure configurations: they follow from the fact that unaccusative verbs have a single direct internal argument and no external argument, while unergative verbs have a single external argument and no internal argument. The phenomena in question could not be captured at a lexical-conceptual level of representation because, in Rappaport and Levin's words:

> "Due to the many-to-one nature of the LCS (Lexical-Conceptual Structure) to PAS (Predicate Argument Structure) mapping, there is probably no single semantic characterization that will distinguish all external arguments from all internal ones...Although it appears that a verb's membership in either of these classes is more often than not predictable from its meaning, it is difficult to give a unified semantic characterization of either class. For example, the class of unergative verbs includes verbs of manner of motion, communication, bodily processes, gestures and sign, and involuntary emission of stimuli, while the class of unaccusative verbs includes verbs of inherently directed motion, change of location, change of state, and appearance and existence. Thus, restating certain generalizations in terms of lexical-semantic representations, while possible, is not very revealing."
> (Rappaport and Levin 1988: 35)

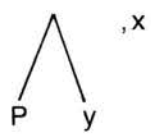
We take this conclusion to be only partially adequate. While we wish to maintain that a level of lexical-syntactic representation is necessary, we will show in the final section

of this chapter that an account of the development of auxiliaries requires reference to members of subclasses of both unergative and unaccusative verbs.

3.9.3 Irregular unaccusatives and marked linking rules: Zubizarreta (1987)

The theory of argument structure developed in Zubizarreta (1987) is based on "a formally richer and more articulated lexical representation" (p.) which distinguishes between the two levels of *lexico-semantic structure* (S-R) and *lexico-syntactic structure* (L-R).[38] S-R represents the grammatically relevant aspects of meaning and consists of structured predicate-argument relations, which represent the selectional properties of predicates. The notions of external and internal arguments are redefined in terms of *scope* of the predicate: internal arguments are within the scope of the predicate, whereas external arguments are outside the scope of the predicate.

(136) a. Transitive



hit y ; x

b. Unergative

P, x work, x

c. Unaccusative

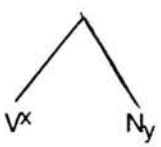
arrive y

As in other theories, thematic roles do not have a place in S-R (with the only exception of *agent* which has the function of distinguishing two types of external argument: +agentive and -agentive).[39] Zubizarreta emphasizes that excluding thematic roles from the set of grammatically relevant features does not entail that there are no substantive semantic notions at work in the grammar. She in fact posits

that the notions of STATE and CAUSATION are necessary in order to formulate some lexical processes that affect meaning.

L-R is the level that mediates the mapping of arguments of verbs and adjectives from S-R to the syntactic level.[40] The mapping consists of three Linking Rules: the Rule of Projection, which requires that internal argument variables be governed by the head of the predicate; the Core Linking Rule and the Default Linking Rule. The latter two rules form open L-predicates. The Core Linking Rule links the external argument variable of a predicate to the head of the L-predicate; the Default Linking Rule links the internal argument variable to the head of the L-predicate. The output of linking rules ("lexical frames") for major verb types is illustrated in (137):

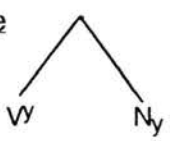
(137) a. Transitive



b. Unergative

V, x

c. Unaccusative



In (137a) and (137b) the Core Linking Rule applies, whereas (137c) results from the application of the Default Rule because unaccusative verbs have no external argument variable.[41]

Zubizarreta argues that there are two types of lexical processes: (a) those that affect predicate-argument relations, and therefore cause a change in meaning, and (b) those that affect the mapping of arguments onto the syntax but not the predicate-argument relations, and therefore leave the meaning unchanged. The former type of process take place at S-R, whereas the latter are naturally described at L-R.

Among the processes that applies at S-R, the process termed 'anti-causativization' is of particular relevance for our purposes: the term refers to the derivation of unaccusatives ("noncausative monadic verbs") from transitive ("causative") verbs. Anti-causativization is brought about by the deletion of the CAUSE feature[42]; the

absence of a 'causer' has as a side effect the deletion of the external argument. The process is illustrated in (138):

(138) sink (1): S-R: <<CAUSE *sink* > y>, x

sink (2): S-R: <*sink* y>

Zubizarreta is concerned with the fact (discussed in section 3.5 above) that while Italian anti-causatives display systematic unaccusative behaviour vis-à-vis the usual diagnostics (selection of ESSERE, NE-cliticization), French anti-causatives divide into two groups: those that have the clitic *se* attached to them satisfy these diagnostics but 'bare' verbs (i.e. those in non-reflexive form) do not. As we saw in section 3.5, this anomalous behaviour is accounted for by Burzio's configurational hierarchy for ESSERE/ETRE selection: Zubizarreta's anti-causatives correspond to the class of paired unaccusatives with a transitive alternant, which are predicted to be peripheral , and therefore variable, in French. What is interesting in Zubizarreta's treatment of these verbs is her attempt to explain their irregularity in terms of a marked mapping between the lexico-semantic and the lexico-syntactic levels:

> "The internal argument, instead of being projected onto the position governed by the head of the lexical frame (by the Rule of Projection) and then linked to head of the lexical frame (by the Default Linking Rule), is directly linked to the head of the lexical frame....It is clear that in French, unlike Italian, the clitic *se* attached to a class of anti-causative verbs serves as a morphological marker to distinguish the anti-causatives that map their S-R onto L-R via the regular linking rules from those that contain an irregular S-R/L-R mapping."
> (Zubizarreta 1987: 90)

This irregular linking is shown in (139):

(139) a. se casser: S-R: <*casser* x>

L-R:

b. couler: S-R: <*couler* x> (irregular linking)

L-R: $\dot{V}^x$

The crucial consequence of this analysis is that if the two types of unaccusative verbs have two distinct L-Rs, then they will have different syntactic structures as well: *se casser* will have the syntactic structure of unaccusatives, while *couler* will be syntactically comparable to unergatives. This conclusion converges with the view expressed in other studies (e.g. Labelle 1990) that French anti-causatives, although they appear to be thematically unaccusative, are in reality unergative. We shall return to this point in section 3.10.

3.9.4 Combining thematic and aspectual relations

Grimshaw's (1990) theory of argument structure, like Rappaport and Levin's, assumes that argument structure is a structured representation, where arguments are relational notions. Specifically, she argues that arguments have prominence relations. The prominence relations are determined by the interaction of the thematic properties of the predicate (defined by the Jackendoff's (1972) Thematic Hierarchy), and by its aspectual properties. These will be now examined in turn.

Thematic prominence is determined by the following scheme, where Agent is the most prominent and Theme the least prominent argument:

(140) (Agent (Experiencer (Goal/Source/Location (Theme))))

For agentive verbs, such as transitives and unergatives, the Agent is always the most prominent argument at the level of argument structure and also the the syntactically most prominent argument, the subject.

(141) a. EAT (x (y)) b. WORK (x)
Agent Theme Agent

It is not always the case that prominence at the level of argument structure and syntactic prominence coincide in the same argument. Some of the so-called psychological-agentive verbs, like *frighten* , are an example of mismatch, because the Theme (the lowest argument) is realized as subject and the Experiencer is realized as direct object. Other psychological verbs like *fear* do not exhibit this mismatch, i.e. the Experiencer becomes a subject and the Theme an object. Both *frighten* and *fear*, however, have an identical thematic structure, which does not reflect their syntactic differences:

(142) a. FEAR/FRIGHTEN (x (y))
Experiencer Theme

Obviously, the syntactic realization of an argument as subject is not dependent on thematic prominence. Grimshaw argues that in order to understand why the Theme of *fear* is realized as direct object and the Theme of *frighten* is realized as subject one needs to consider a second semantic dimension: aspectual prominence.

With respect to this dimension, the difference between *fear* and *frighten* , as a first approximation, is that the latter has a causative component associated with it, whereas the former does not. Although the arguments of the two verbs are the same at the thematic level, they differ because the Theme is a cause in *frighten* but not in *fear*. The causal status of an argument determines whether it is realized as a subject or not: if it is a cause, then it must be realized as a subject.

Thus, the causal status of an argument identifies a dimension of prominence relations distinct and autonomous from the thematic dimension. The causal structure of a predicate defines an aspectual hierarchy, in which the cause argument is the most prominent. Both the thematic and the aspectual hierarchies impose their own prominence relations on the set of grammatical arguments projected by a predicate. So, in (143), the predicate *break* has to be considered in terms of both dimensions:

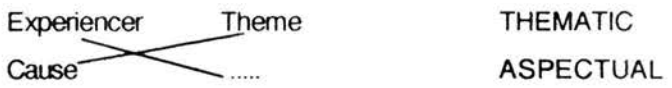
(143) a. The girl broke the window

b. BREAK (x (y))

Agent	Patient	THEMATIC
Cause	.....	ASPECTUAL

The argument in the subject position of *break* is more prominent than the other in terms of both dimensions, since it is a cause and a thematic agent. In the case of *frighten*, on the other hand, the second element in the aspectual dimension is associated with the first element in the causal dimension, and vice versa:

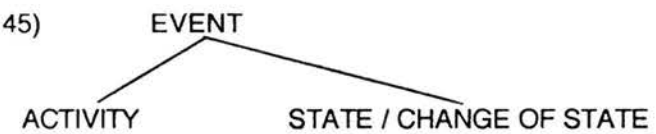
(144) FRIGHTEN (x (y))



The special character of this class of verbs is therefore given by the misalignment of the thematic and the aspectual dimensions: the subject is most prominent in the causal hierarchy but not in the thematic hierarchy. This is what distinguishes the *frighten* class not only from the *fear* class, but also from transitive agentive verbs (like *arrest*) and unergative verbs (like *work*), for which the alignment of the two dimensions is always realized.

These latter verbs, however, have no causative component in their meaning, which suggests that the notion of cause is too narrow. Grimshaw suggests that, in more general terms, all verbs may be conceived of as part of an event structure which, combined with other elements of the clause provides an event structure for the whole sentence. This event structure can be represented as in (145), which shows a Vendler-Dowty accomplishment: 'activity' and 'state' are sub-events that characterize the action described by the verb:

(145)



The causal argument is always associated with the first sub-event, which is causally related to the second sub-event. The basic generalization is that an argument which participates in the first sub-event is more prominent than an argument that participates in the second sub-event. Since a cause is always part of the first sub-event, it is always more prominent than the argument corresponding to the entity which undergoes a change of state. Consequently, agents will always be subjects not because of their thematic properties alone, but because of their combined thematic and aspectual prominence.

3.9.4.1 Subjects and external arguments

The two-dimensional approach to prominence makes it possible to define the notion of 'external argument' in an alternative way.[43] External arguments are arguments that are maximally prominent along both the thematic and the aspectual dimensions. Thus, a verb lacks an external argument if no single argument satisfies the condition of maximal prominence along the two dimensions. This is the case of the *frighten* class, where the Experiencer is maximally prominent thematically but not aspectually. Verbs like *frighten*, however, have a d-structure subject (the cause of the action): this shows that, while all external arguments are also subjects, not all subjects are external arguments. The notion of external argument is defined in terms of argument structure representation, whereas the notion of subject is defined in configurational terms.

The other major class of verbs lacking an external argument is the class of unaccusative verbs. Unaccusatives have no external argument, because, according to the maximal prominence criterion, their single argument fails to reach maximal prominence either in the thematic or in the aspectual dimension, or both. It could be argued that unpaired unaccusative verbs have a single argument (a Theme) that ought to be maximally prominent for the simple reason that there is no other argument to compete for this assignment. Themes, however, never count as maximally prominent, and are therefore never realized as external arguments, because they belong to particular aspectual classes: they express states or changes of state. The event structure of an unaccusative verb, corresponds to the second sub-event of an accomplishment (as in (146c), whereas the event structure of an unergative verb corresponds to the first sub-event (as in (146b):

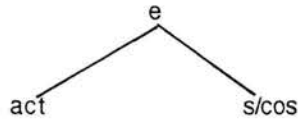
(146) a. accomplishment
[[activity] [state / change of state]]

b. unergative
[[activity]]

c. unaccusative
[state / change of state]

The argument of an unaccusative verb does not participate in the first sub-event, and therefore lacks aspectual prominence. As such, it does not meet the criterion for externality.

In general terms, this line of argument leads Grimshaw to conclude that:

> "...This result will follow if what determines the aspectual hierarchy is an event-structure template which is fixed for all predicates rather than being projected from the lexical-semantic representation of the individual predicate. The aspectual dimension, then, is a projection of an abstract event structure (e), which always includes two subparts, an activity (act) and a state or change of state (s/cos):
>
> 
>
>
> The event template determines prominence, assigning the maximally prominent position in the aspectual dimension only to arguments participating in the first subevent, regardless of the actual lexical semantic representation of the predicate. If all events are constrained by this template, activities will always fit the first slot in the template, and an existential state will always fit the second slot. Thus a single argument of an unaccusative will never count as maximally prominent and will never qualify as external."
> (Grimshaw 1990: 40)

This approach has the advantage of exploiting the syntactic relevance of lexical semantic characterizations, unlike purely semantic treatments of unaccusativity such as van Valin's. Grimshaw points out (like Rappaport and Levin) that a purely syntactic level of representation is necessary to account for phenomena associated with unaccusativity that are unlikely to reduce to the argument structure status of arguments. These phenomena can be best explained by referring to d-structure objects, rather than internal arguments. NE-cliticization in Italian is an example of syntactic behaviour that is not satisfactorily accounted for in terms of thematic or aspectual hierarchies:

> "It seems considerably more likely that the system of government or related syntactic notions lie behind the phenomenon, as is usually assumed. Such an explanation implies that the argument of an unaccusative must be governed and hence must be an object, not a subject."
> (Grimshaw 1990: 42)

These considerations bring us back to the possibility (already mentioned in section 3.8) of regarding the various phenomena related to unaccusativity in terms of their sensitivity to a semantic characterization: some types of grammatical behaviour could receive both a syntactic and a semantic definition, while other types would be resistant to a semantic definition. The latter are derived as a consequence of an argument being in direct object position, while the former are derived from the thematic, of conceptual status of an argument. This possibility will be explored in the next section.

3.10 Lexical-semantic hierarchies

We noted earlier that Italian has preserved the semantic complementarity of the Late Latin system, which was governed by the principle that verbs with an agentive or affecting subject selected HABERE and verbs with a patient or affected subject selected ESSE (Vincent, 1982; Tuttle, 1986). In modern Italian, there still is a clear and consistent correlation between the choice of ESSERE and non-agent thematic properties of the subject of the verb. The broad concept of THEME as 'affected entity' encompasses the definitions of such properties that have been offered in the literature: neutral or patient (Vincent, 1982), theme (in both the literal and the metaphorical sense of 'object which moves') (Keenan, 1987), undergoer (Van Valin, 1990), affected actor (Centineo, 1986), argument of a state or of a change of state (Parisi, 1978; Grimshaw 1990). We want to suggest that a greater differentiation within the THEME construct allows an explanatory account of both diachronic and syncronic variation in the domain of auxiliary selection.
For this purpose, we will provide evidence from historical language change and from variability in contemporary auxiliary usage in French and Italian.

3.10.1 The erosion of ESSE-reflexes in Romance

We mentioned in section 3.8.2 that derivations of Latin ESSE as auxiliary have their last bastions in the auxiliary systems of Italian and French. In the majority of Romance languages, the domain that used to be of ESSE was gradually eroded, and then completely taken over, by reflexes of HABERE.

Benzing (1931) provides a painstakingly detailed, and still unrivalled, analysis of the disappearance of SER as an auxiliary in Spanish from the 13th till the late16th century. Using a three-way classification of intransitive verbs into verbs of movement (Verben der Bewegung), verbs of change of state (Verben des Entstehens und Vergehens: literally 'coming into and going out of existence'), verbs of state (Verben der Ruhe), he shows that SER was replaced by HABER first with verbs of state, then with verbs of change of state, and last with verbs of movement:

> "... *ser* [stirbt] zuerst bei den Verben der Ruhe aus, dann bei den nur selten gebrauchten Verben des Bewegung wie *exir, arribar, avenir*, auch *tornar*. Die Verben des Entstehens und Vergehens nehmen mehr oder weniger eine Mittelstellung ein. Während *nacer* und *morir*, *ir* und *pasar* noch über das 16. Jahrhundert hinaus *ser* als Hilfsverb gebrauchen, kann man doch sagen, dass mit dem Ende des 16, Jahrhundert *ser* als Hilfsverb von *haber* verdrängt ist."
> (Benzing 1931: 459).

The most frequently used verbs of movement were therefore the most resistant to penetration by HABER, while stative verbs were the most vulnerable to it. Relics of the old auxiliary system are present even in modern Spanish, given that expressions like *es ido* 'he is gone', and *es venido el momento* 'the moment has come' are not completely unacceptable to native Spanish speakers.
An analogous evolution of the ESSE-reflexes can be retraced in Catalan (Tuttle 1986: 264-265), and Provençale (Savinian 1882). In these languages, ESSE-reflexes were used more widely centuries ago, with the same types of verbs, and has survived in the form of isolated occurrences with verbs of movement.[44]

In Old French, ETRE was used with a much wider range of verbs than in modern French, including verbs such as *voler, couler, courir,* and *être* itself (Gougenheim 1951: 122-123).[45] Of the handful of verbs that have survived the penetration of

AVOIR until today, most are verbs of movement, followed by verbs of change of state (Grevisse 1980), and by a few verbs denoting continuation of state.

3.10.2 Auxiliary leveling in Canadian French

The process of erosion of ETRE has progressed even further in varieties of French separated from Standard European French by spatial, cultural, or political distance. One of these varieties is Canadian French, where ETRE is still consistently used only with a much smaller number of verbs than in Standard French.

Sankoff and Thibault (1977) describe the variable auxiliary use in spoken French in Québec. They analysed a corpus of 120 taped conversations in terms of probability of use of AVOIR instead of ETRE with a set of verbs that can be classified into the two classes of [+complete] and [-complete] action. Their findings suggest the following hierarchy of verbs, ordered from least to most likely to be employed with AVOIR:

(147)

aller
revenir
venir
entrer
arriver

partir
retourner
descendre
monter
sortit
rentrer
tomber
déménager
passer

changer

rester
demeurer

- AVOIR
↓
+ AVOIR

The verbs in (147) are grouped, according to Sankoff and Thibault, in terms of the possibility of using their past participles as stative adjectives . Their hypothesis is that use of AVOIR as an auxiliary goes hand in hand with the [+complete] semantic trait of a verb, and therefore, with the low likelihood of an adjectival use of its past participle. While this seems plausible, other considerations are made possible by this study. The verbs in (147) seem to be ordered with respect to two

characteristics: one is a dimension whose poles can be defined as 'dynamic vs static', so that the verbs at the top of the list are the ones that most properly denote change of location, while the verbs at the bottom of the list denote continuation of an existing state; the other is the distinction between paired and unpaired unaccusatives, so that the verbs in the group at the top of the list are all unpaired, whereas all the others (with the exception of *rester* and *demeurer*) have transitive alternants. What the list in (147) suggests is that the verbs that are most impervious to AVOIR are unpaired verbs denoting change of location.

The results of Sankoff and Thibault's study are supported by Canale, Mougeon and Bélanger (1978), who looked at variation in the frequency of occurrence of AVOIR based on the spontaneous speech of Franco-Ontarian students in Grades 2, 5, 9, 10 and 12 at French schools in Ontario. The findings show a tendency to use AVOIR instead of ETRE according to the following hierarchy of verbs of change of location (the only exception is *rester,* which denotes continuation of an existing state):

(148)

Again, unpaired change-of-location verbs are found to be more impervious to the ETRE-->AVOIR replacement than paired verbs and continuation-of-state verbs. Secondary school students tend to use AVOIR less frequently with the verbs in (148) than do the elementary school students: this indicates the still relatively strong influence of literacy and metalinguistic knowledge in counterbalancing the progressive loss of ETRE as an auxiliary.

3.10.3 Changes in auxiliary usage in modern Italian

Finally, we turn to variable auxiliary use in Italian. Despite its conservatism, Italian has not been immune to the pan-Romance tendency towards the elimination of ESSE-reflexes as auxiliaries. Although it can be said that the process, in historical terms, has just begun in Italian, and has not been registered at the normative level, its effects have started to become visible in the colloquial language, generating variation in auxiliary usage with respect to certain types of verbs. In the absence of a systematic study on this phenomenon, it is interesting to assemble the various observations that have been made in a limited number of descriptive studies.

Rohlfs (1969) comments on the general 'promiscuous' use of auxiliaries with verbs like *appartenere* 'belong, *durare* 'last', *fiorire* 'blossom', *vivere* 'live', *sfilare* 'parade', *essere* itself, and reflexive verbs. Furthermore, he describes the marked generalization of AVERE in Southern Italy, where forms such as *ho stato* 'I have been', and *ho rimasto* 'I have remained' are common. In the North-East of Italy AVERE is also largely generalized over ESSERE,particularly with impersonal stative verbs (*piacere*, 'like', *suonare* 'ring'), and 'weather' verbs. In these dialects, however, one can observe a preference for ESSERE as the auxiliary of verbs of movement (see also Tuttle 1986).

Berruto (1987:20) underlines the 'marked development of AVERE' in contemporary Italian usage, and its consistency with the 'pan-Romance tendency'. He reports examples of the extension of AVERE such as the ones in (149):

(149) a. Ancora una volta ho rimasto solo
'Once again I have been left alone'

b. Quando la macchina è nuova e non si ha avuto il tempo di fare applicare...
'When a car is new and one has not had the time to apply...

The sentence in (149a) contains a verb denoting continuation of state, while the sentence in (149b) shows a use of impersonal SI with AVERE.

Jacqmain and Meerts (1983) argue that the so-called 'weather verbs' have gone, in a diachronic perspect, from an almost exclusive use of ESSERE as an auxiliary to the present undifferentiated use of either ESSERE or AVERE.[46] The verb *vivere* 'live' is an example of the 'spectacular advancement of auxiliary AVERE', since it is now widely used with this auxiliary. This preference also applies to compounds of *vivere* such as *convivere* 'live with', *rivivere* 'live again'.
The case of *sopravvivere* 'survive' is emblematic of the growing divergence between a normative attitude, which prescribes ESSERE as the obligatory auxiliary, and examples such as (150), which can be encountered in actual use of the language:

(150) La Comunità ha sopravvissuto al Consiglio Europeo
'The Community has survived the European Council'

All the examples we have mentioned of variable auxiliary use in modern Italian seem to concern unaccusative stative verbs, or verbs denoting continuation of an existing state, but no verbs of change of state. This raises the question of systematicity in development. Is there a path in the evolution of auxiliaries that will eventually lead to the complete loss of HABERE-reflexes in Romance?

3.10.4 The Unaccusative Hierarchy

Our proposal is that a fine-grained analysis of the arguments subsumed under the label 'theme' of unaccusative verbs reveals that in the majority of Romance languages HABERE-reflexes have been diachronically generalized at the expense of ESSE-reflexes *according to a systematic pattern*. This becomes clear if one posits certain distinctions within the range of unaccusative verbs, which differentiate among the types of process undergone by the subject of the verb.

If the conventional representation of the argument structure of unaccusative verbs is 'VERB: <VERB y>, the verb classes along the hierarchy (exemplified below in both Italian and French) can then be represented as the following substructures at the lexical-semantic level:

(151) A hierarchy of unaccusative verbs

(a) Unpaired unaccusatives

ANDARE/ALLER:	< y comes to be at a different LOCATION>
DIVENTARE/DEVENIR,	
SPARIRE/DISPARAITRE:	< y comes to be in a different CONDITION>
DURARE/DURER:	< y continues to be in the same CONDITION>
ESISTERE/EXISTER:	< y is in some CONDITION >

(b) Paired unaccusatives:

AUMENTARE/AUGMENTER:	< y comes to be in a different CONDITION>
	(with alternant AUMENTARE: <aumentare y, x >)
CORRERE/COURIR:	< y comes to be at a different LOCATION >
	(with alternant CORRERE: < correre x >)

The hierarchy in (151) distinguishes between paired and unpaired unaccusative verbs, and orders unpaired unaccusatives according to their semantic status with respect to the dimension 'DYNAMIC vs STATIC'. The hierarchy embodies the hypothesis that the notion of dynamic change, whose most concrete manifestation is change of location, is at the root of unaccusativity, and identifies verbs of directed motion as core cases for ESSERE/ETRE-selection. The unpaired verbs in (151a) seem to be also ordered with respect to the dimension 'CONCRETE vs ABSTRACT': core verb types denote concrete change of location (i.e. movement in space from A to B), while the further a verb type is from the core, the more abstract its meaning is.

The hierarchy provides a key to an overall interpretation of the diachronic evolution of auxiliaries in Romance. Among the unpaired unaccusatives, 'change of location' verbs have been more impervious to the expansion of HABERE-forms than 'change of condition' verbs, which in turn have been less open to change than 'abstract existential verbs'. These are followed by paired verbs that, in addition to the unaccusative version, have either a transitive or an unergative alternant, and which are therefore also vulnerable to analogic levelling because of the auxiliary selected by their lexical counterparts.

We can therefore say that the process of change has been spreading from the periphery towards the core of a hierarchy of unpaired unaccusative verbs based

on the dimensions of concreteness/abstractness and movement/staticity. The hierarchy is illustrated in Table 3.8 for Italian and French:

TABLE 3.8 The Unaccusative Hierarchy in Italian and French

VERB TYPE	SEMANTIC DIMENSION	DIACHRONIC DIMENSION	FRENCH	ITALIAN
CHANGE OF LOCATION (andare, venire.., aller, venir...)	CONCRETE, MOVEMENT	-open to *habere*-reflexes	ETRE	ESSERE
CHANGE OF CONDITION (diventare, sparire... devenir, disparaître...)			ETRE/AVOIR	ESSERE
CONTINUATION OF A CONDITION (sopravvivere, durare.. survivre, durer)			AVOIR[47]	ESSERE
EXISTENCE OF A CONDITION (essere, esistere, sembrare... être, exister, sembler)	ABSTRACT, STATICITY		AVOIR	ESSERE
+TRANSITIVE ALTERNANT (aumentare, migliorare... augmenter, ameliorer)			AVOIR	ESSERE
+UNERGATIVE ALTERNANT (correre, rotolare.. (courir, rouler)		+open to *habere*-reflexes	AVOIR	ESSERE

The hierarchy in Table 3.8 makes clear the integration between the syntactic and the semantic components of unaccusativity. The unaccusative hierarchy intersects Burzio's configurational hierarchy in Table 3.1, repeated in Table 3.9:

TABLE 3.9: Configurational hierarchy for ESSERE-and ETRE-selection in Italian and French (adapted from Burzio, 1986: 140)

	ITALIAN	FRENCH	
a. NP cl-V	core	core	(unaccusatives in reflexive form)
b. NP V NP	core	periphery	(unaccusatives in non-reflexive form)
c. NP V [$_S$ NP...]	periphery	-------	(Raising and restructuring constructions)

Within configuration (b) one finds systematicity, not random variation: while all the Italian verbs along the hierarchy select ESSERE, the verbs selecting ETRE in French are restricted to the 'change of location' and the 'change of condition' categories. Furthermore, it is not coincidental that most unaccusative 'mismatches' (see sections 3.5 and 3.9.3 above) are found within the category of paired unaccusatives with a transitive alternant, which is relatively low in the hierarchy.

Furthermore, the hierarchy provides a framework for the interpretation of synchronic variation in Canadian French, where the ETRE-->AVOIR change has progressed further towards the top of the hierarchy than in European French, sparing only 'change of location' verbs, and in Italian, where all the attested cases of variation in auxiliary usage fall within peripheral categories.

3.10.5 Heterogeneity of unergative verbs

If the class of unaccusative verbs is semantically heterogeneous, and the notion of change of location is at the core of unaccusativity, can one maintain that the class of unergative verbs is homogeneous?

As we saw earlier, most semantic theories of auxiliary selection agree in identifying 'agentive activity' with no resulting change of state as the thematic-aspectual core of unergativity. It is clear that there are differences among unergative verbs as to the extent they fit the core. Non-motional unergative verbs (*lavorare, parlare, telefonare*, etc) denote pure activities which do not imply a change of location. Motional unergatives (*nuotare, camminare, viaggiare*, etc), like all verbs of motion, have a change-of-location component in their semantic which places them further away from the core. Finally, motional unergative verbs with an unaccusative alternant (*correre, volare, saltare,* etc) are openly ambiguous and can only be

disambiguated contextually by the presence or absence of a locative expression indicating the direction of motion, or by the presence or absence of an adverbial phrase specifying the manner of motion.

On these bases, it is plausible to draw distinctions among the following sub-classes of unergative verbs, ordered with respect to their consistency with the core:

(a) NON-MOTIONAL ACTIVITY (*dormire* 'sleep', *parlare* 'talk, *lavorare* 'work', etc)
(b) MOTIONAL ACTIVITY (*camminare* 'walk', *nuotare* 'swim', *viaggiare* 'travel', etc)
(c) MOTIONAL ACTIVITY WITH UNACCUSATIVE ALTERNANT (*correre* 'run', *saltare* 'jump', *rotolare* 'roll', etc).

If the conventional representation of the argument structure of unergative verbs is taken to be VERB: <VERB x>, the unergative subclasses can be represented (like the unaccusative subclasses in (151)) as sub-structures at the lexical-semantic level:

(152) A hierarchy for unergative verbs

DORMIRE/DORMIR	<x is engaged in an activity>
NUOTARE/NAGER	<x is engaged in an activity involving a change of location>
CORRERE/COURIR	<x is engaged in an activity involving a change of location> (with alternant CORRERE: <y comes to be at a different LOCATION>)

It should be pointed out right away that the hierarchy of unergative verbs proposed in (152) is not as strongly supported by evidence as the unaccusative hierarchy. While the argument 'theme' is widely recognized as being a multifaceted construct, there is no comparable evidence that the argument 'agent' is so heterogeneous in nature. Furthermore, there is no evidence from language change concerning auxiliary choice with unergative verbs. It should be therefore emphasized that, at this stage, the unergative hierarchy is a natural extension of the theoretical arguments employed to posit the latter , and does not have independent justification. Nevertheless, to the extent that the distictions among unaccusative verbs turn out to be empirically reflected by systematic variability in native linguistic intuitions, it is plausible to hypothesize that those underlying unergative verbs will also do so:

confirming evidence would strengthen the notion of relative linguistic acceptability, and the importance of the lexical-semantic level of representation.

3.10.6 Notes on the syntactic realization of lexical-semantic verb types

We have not addressed the question of how the sub-structures in (151) and (152), representing different types of unaccusative and unergative verbs at the lexical-semantic level, are mapped onto variables in argument structure. The mapping mode is crucial because, as we suggested in 3.9.3, it determines the syntactic status of the unaccusative and unergative verb types. If we adopt Rappaport and Levin's approach, we can posit two very general linking rules of the same form as (134) that would access the sub-structures in (151) and (152) and realize the x variables in (151) as external arguments and the y variables in (152) as direct internal arguments in argument structure:

(153) When the lexical-conceptual structure of a verb includes one of the substructures in (151), link the variable represented by y to the direct argument variable in the verb's argument structure.

(154) When the lexical-conceptual structure of a verb includes one of the substructures in (152), link the variable represented by x to the external argument variable in the verb's argument structure.

This linking rule would ensure that all lexical-semantic types of unaccusative and unergative participate in the same syntactic phenomena as a class, which implies that the syntactic properties of unaccusativity/unergativity are blind to lexical-semantic characterizations. However, our overview of auxiliary selection has shown that , at least in French, some unaccusative verb types satisfy syntactic diagnostics of unaccusativity categorically, while other unaccusative types satisfy them variably: such syntactic variation raises the possibility of irregular linking rules which, as we saw earlier, is related to the full membership of 'variable' verbs to the unaccusative class. Furthermore, the question of the irregular syntactic behaviour of verbs has to be considered in the wider context of diachronic change. It was pointed out in section 3.8.2 that the evolution of auxiliaries in Romance has been following an identifiable semantic path, but that the details of the syntactic reanalysis that leads from a double-auxiliary system to a single-auxiliary system

have not yet been investigated. One could speculate, not inconsistently with some of the current research on language change, that changes at the lexical-semantic level create the conditions for syntactic change, and that irregular linking rules could be regarded as one of the early manifestations of syntactic reanalysis in the auxiliary system.[48] In this thesis, however, we shall not try to substantiate such speculation, nor shall we address in any detail the question of the mapping of lexical-semantic representations onto syntactic representations.

We take the position here that there are phenomena in Italian (such as the selection of perfective auxiliaries) that are sensitive to both syntactic and semantic characterizations, and phenomena (such as NE-cliticization and restructuring) that are sensitive only to syntactic characterizations. We predict that in Italian, given the syntactic consistency of the auxiliary system, the verb hierarchies in (151) and (152) affect the former but not the latter: this prediction would be denied if it turned out that syntactic phenomena ARE sensitive to the position of verbs along the hierarchy. Our hypotheses are articulated in detail in the next chapter.

CHAPTER 4
THE EMPIRICAL HYPOTHESES

Chapters 1, 2, and 3 have provided the theoretical background to this study. The main arguments put forward can be summarized under the following three headings: psychological, methodological, and linguistic.

4.1 Psychological arguments

(a) Linguistic acceptability is a relative, not an absolute, property.
(b) Indeterminacy is a natural characteristic of languages. Conceptually, it can be defined as the indefiniteness of grammaticality status of certain constructions in the speaker's mentally represented grammar. Indeterminacy leads to variation and inconsistency in the speaker's acceptability judgments. An empirical measure of indeterminacy is therefore the degree of variation/inconsistency in judgments of linguistic acceptability.
(c) Speakers' knowledge of language may be represented as consisting of an indefinite number of acceptability hierarchies ranging from a determinate core to an indeterminate periphery.
(d) Indeterminacy pervades the interlanguage grammars of adult second language learners at all levels.

4.2 Methodological arguments

(a) Ranking methods for the elicitation of acceptability judgments are more adequate than rating methods for the purpose of capturing variation and indeterminacy because they require relative rather than absolute judgments.
(b) Among ranking methods, those that produce ordinal scales do not allow an interpretation of the distance among points on the scale, and do not readily permit parametric statistical analyses.
(c) If linguistic acceptability is regarded as a non-metric continuum, and if acceptability judgments are seen as psychophysical perception without a corresponding physical measure to serve as a basis for comparison, magnitude estimation is a potentially powerful method for the investigation of variation in acceptability because it makes it possible to measure it directly on an interval scale.

4.3 Linguistic arguments

(a) Auxiliary selection with intransitive verbs in Italian (and in other languages) is governed by both syntactic and semantic constraints.

(b) The two major types of intransitive verbs - unaccusative and unergative - have different syntactic representations, as the Unaccusative Hypothesis maintains: the surface subject of unaccusatives is a d-structure direct object, while the surface subject of unergatives is a d-structure subject. These syntactic differences are responsible for the systematically different behaviour of unaccusative and unergative verbs vis-à-vis a number of syntactic diagnostics. Optional restructuring rules in Italian are also dependent on the syntactic representation of unaccusativity embodied by the Unaccusative Hypothesis.

(c) Italian intransitives are characterized by a high degree of syntactic coherence. Auxiliary selection and NE-cliticization, as two of the syntactic manifestations of the unaccusative/unergative distinction, are almost entirely predictable. In contrast, French is syntactically less coherent: auxiliary selection and EN-cliticization do not unambiguously determine the membership of intransitive verbs to the unaccusative or the unergative class. The selection of ESSERE/ETRE in Italian and French is parameterized and can be represented by a syntactic hierarchy based on the notion of government, ranging from core to peripheral configurations. Unaccusativity in English, on the other hand, is more easily characterizable in semantic, rather than syntactic terms.

(d) Class membership for the unaccusative or the unergative classes may also be determined by semantic principles. In terms of thematic roles, the internal argument of unaccusative verbs is a theme and the external argument of unergative verbs is an agent. In terms of aspectual relations, unaccusative verbs predicate a state or a change of state of their argument, whereas unergative verbs predicate an activity. These generalizations have validity as statistical tendencies, not as universals.

(e) Some of the phenomena related to the unaccusative/unergative distinction (e.g. NE-cliticization and restructuring rules) are sensitive only to syntactic representations; others (e.g. the selection of perfective auxiliaries) are sensitive to both syntactic and semantic representations.

(f) It is proposed in the present study that further lexical-semantic distinctions within the class of unaccusative verbs are necessary in order to shed light on the path of diachronic evolution of auxiliaries in Romance languages, and on the variation in auxiliary usage found in modern Italian and modern French. The broad

concept of 'theme' can be analysed in terms of the type of process affecting the subject of the verb, with respect to the dynamic vs static dimension. The result is an acceptability hierarchy for unpaired unaccusatives ranging from 'change-of-location verbs at the core to 'existence-of-state' verbs at the periphery. The extension of the scope of HABERE-reflexes at the expense of ESSE-reflexes appears to have been affecting Romance languages by starting at the periphery of the hierarchy and gradually spreading to the core.
It is also suggested that the class of unergative verbs is not semantically homogeneous and can be subdivided into sub-types according to whether they do or do not denote activities which involve an implicit change of location.

4.4 Hypotheses

The theoretical concepts outlined above motivate the hypotheses that we want to submit to empirical testing. These hypotheses are of two kinds. The first kind are based on the methodological arguments, while the second kind refer to the linguistic and psychological arguments. Both methodological and linguistic experimental hypotheses are stated below: each hypothesis is followed by the enunciation of the null hypothesis (i.e. the 'no effect' hypothesis that would contradict our predictions, and that we hope to reject).

From a methodological point of view, we hypothesize that:

(a) Magnitude estimation will be a viable method for the collection of linguistic acceptability judgments from both native and non-native speakers.
(b) Magnitude estimation will produce results as least as informative as those produced by card-sorting.
(c) Magnitude estimation will be more discriminating than card-sorting in distinguishing degrees of acceptability.
(d) Magnitude estimation will reveal meaningful and interpretable patterns of acceptability in judgments that will not be captured by card-sorting.

[Null hypothesis: magnitude estimation does not work with linguistic acceptability judgments. It produces less discriminating and less consistent results than card-sorting. It reveals uninterpretable patterns of acceptability].

From the linguistic point of view, we predict that:

(a) the different grammatical phenomena related to auxiliary selection should be associated with different degrees of determinacy in the linguistic intuitions of native Italian speakers.

(b) these different phenomena should also be associated with different degrees of learnability in the acquisition of Italian auxiliaries by native speakers of other languages.

The two hypotheses can be articulated in detail as follows:

(a) the DETERMINACY hypothesis:

i. aspects of auxiliary selection that have a semantic basis will elicit systematically variable responses at the level of native acceptability judgments, whereas aspects that have a purely syntactic origin will elicit uniform responses (i.e. categorical acceptances or categorical rejections). [Null hypothesis: there is no difference in terms of variation/uniformity between native judgments on semantically-based aspects and judgments on syntactically-based aspects of auxiliary selection].

ii. specifically, auxiliary choice in the present perfect will be conditioned by the lexical-semantic distinctions underlying the proposed acceptability hierarchies for unergative and unaccusative verbs. These distinctions correspond to an acceptability scale in native intuitions: core verb types are more determinate than peripheral verb types. Therefore, native speakers will not judge all verb types within the unaccusative and unergative classes in the same way. The degree of consistency should be a function of the position of a verb category on the hierarchy it belongs to: judgments on core categories will be more determinate and consistent than judgments on peripheral categories.
[Null hypothesis: judgments on auxiliary choice with unaccusative and unergative verbs have the same degree of determinacy, irrespective of lexical-semantic distinctions within the two verb classes].

iii. syntactically-derived phenomena like NE-cliticization and Restructuring will elicit uniform judgments from native speakers, because they should not be sensitive to lexical-semantic characterizations. This means that native speakers will accept all unaccusative verbs presented with NE-cliticization and ESSERE and reject all unaccusative verbs presented with NE-

cliticization and AVERE, regardless of their lexical-semantic class; they will also reject NE-cliticization with unergative verbs, regardless of auxiliary and lexical-semantic verb type. Judgments on NE-cliticization, however, are predicted to have an overall higher degree of determinacy than auxiliary selection under restructuring, since the latter is peripheral in terms of the syntactic hierarchy based on government.
[Null hypothesis: NE-cliticization and Restructuring are judged differently depending on the type of verb they are presented with].

(b) the LEARNABILITY hypothesis:

i. aspects of auxiliary selection that have a semantic basis should be more learnable in Italian as a foreign language than aspects that have a purely syntactic origin, possibly because of the reduced availability of Universal Grammar in second language acquisition. Overall, non-native acceptability judgments on semantic aspects will be more determinate than non-native judgments on syntactic aspects.
[Null hypothesis: there is no difference between semantically-based and syntactically-based aspects of auxiliary selection in the judgments of non-native speakers: they are equally determinate/indeterminate]

ii. lexical-semantic hierarchies of unergative and unaccusative verbs should affect the order of acquisition of perfective auxiliary selection with particular verb categories, i.e. the degree of difficulty with which auxiliaries are acquired. The order of acquisition should be related to the position of a given verb category along its hierarchy: core categories should be learned earlier than peripheral categories. Non-native intuitions should develop in gradual approximation to native intuitions.
[Null hypothesis: there are no differences within the class of unaccusative verbs, nor within the class of unergative verbs, with respect to the degree of difficulty in acquisition].

iii. The development of linguistic intuitions about syntactically-based phenomena like NE-cliticization and restructuring should not depend on lexical-semantic verb types and should not follow a path of gradual and steady approximation to native values. Acquisition - if it happens at all - should be discontinuous.

[Null hypothesis: the path of development of syntactically-based phenomena is gradual and continuous from the beginning to the near-native proficiency level. NE-cliticization is acquired first for core verb types and last for peripheral verb types]

(iv) The native language background of learners should determine whether and to what extent the syntactically-cased phenomena are acquired. French is predicted to be a more favourable starting point for learning Italian than English is: this is because the French system still has syntactic reflexes of the unaccusative/unergative distinction which are in parametric variation with the corresponding Italian ones (i.e. ETRE-assignment and EN-cliticization), whereas English does not instantiate any of these.
[Null hypothesis: there is no difference between French and English learners of Italian with respect to the acquisition of syntactically-based properties of auxiliary selection].

v. the differences between the auxiliary systems of Italian and French (in terms of internal consistency and predictability) create learning asymmetries, such that it is easier for French learners of Italian to acquire the more coherent Italian system than for Italian learners to acquire the less coherent French system.
[Null hypothesis: there is no difference in terms of degree of difficulty between the acquisition of French auxiliaries by Italian speakers and the acquisition of Italian auxiliaries by French speakers].

The empirical hypotheses presented in this chapter were tested first in a series of pilot studies, and then in larger-scale experiment. The next chapter will be concerned with the pilot studies.

CHAPTER 5
THE PILOT STUDIES

5.0 Introduction

The previous chapters have identified variation in linguistic acceptability as the domain of investigation of this study. Two main research issues have been raised:

(a) the existence of acceptability hierarchies in languages, ranging from determinate core constructions to indeterminate peripheral constructions. Auxiliary selection in Italian has been shown to be a non-uniform grammatical area, where semantic and syntactic factors influence each other in theoretically predictable ways.
(b) the need for an adequate technique for the elicitation of acceptability judgments, which captures systematic variation in acceptability. Magnitude estimation has been singled out as a potentially powerful instrument for this purpose, which combines the advantages of a fine-grained discrimination of acceptability with the applicability of parametric statistics.

As the existence of both acceptability hierarchies and a technique for discovering them were in doubt, both were the subject of pilot studies. These were designed to determine:

(1) the relative validity of Magnitude Estimation as a technique for the elicitation of acceptability judgments with respect to more conventional ranking methods;
(2) the applicability of the technique to native and non-native speakers;
(3) its ability to reveal systematic variation in linguistic intuitions from a determinate core to an indeterminate periphery.

Three pilot studies were conducted, which will henceforth referred to as Pilot 1, Pilot 2 and Pilot 3:

- Pilot 1 looked at variation in the area of auxiliary selection with intransitive verbs, and predicted asymmetric patterns in the acquisition of French auxiliaries by Italian learners and the acquisition of Italian auxiliaries by French learners which could be attributed to the Unaccusative Hierarchy.

- Pilot 2 was also concerned with auxiliary selection and past participle agreement with transitive, unergative and unaccusative verbs; in addition, it focused on

different methods for eliciting acceptability judgments from native and non-native speakers and for capturing the variation in their linguistic intuitions.

- Pilot 3 aimed at detecting systematic variation within a different area of grammar: the aspectual distinction between perfectivity and imperfectivity in Italian. Its principal focus was on empirical methods for the elicitation of acceptability judgments.

Inevitably, each pilot study could address only a subset of the issues involved in this research; furthermore, they were the testing ground for ideas and materials that were - a posteriori - found inadequate and therefore discarded. Despite their limitations, however, the pilot studies were not only useful in suggesting improvements in the experimental materials and the research design, but also instrumental for the precise definition of the research questions addressed by the main study.

5.1 Pilot 1: asymmetries in learnability

Second language acquisition of ESSERE/ETRE in Italian and French

The first pilot study was exclusively concerned with theoretical hypotheses about auxiliary selection, rather than methods of elicitation. The starting point was the difference between Italian and French with respect to auxiliary selection with unaccusative verbs. As was noted in Chapter 3, all Italian unaccusative verbs select ESSERE, but only a handful of French unaccusative verbs select ETRE. French is at a more advanced stage in the diachronic process of replacement of reflexes of Latin ESSE by reflexes of HABERE. This has led to a loss of internal consistency in the French auxiliary system; in contrast, the Italian system has preserved the semantic equilibrium of the Latin system, and its class of unaccusative verbs is syntactically consistent.

The asymmetry between French and Italian auxiliaries allows a prediction pertinent to second language acquisition, namely that despite the superficial similarities between French and Italian auxiliaries Italian learners of French and French learners of Italian would not face the same task in acquiring the auxiliary system of the foreign language. This prediction is theoretically interesting because it can result from two different sets of theoretical assumptions.

The first set of assumptions predicts that, given that the French system for ETRE assignment is considerably less consistent than the Italian system for ESSERE assignment, greater demands are placed on Italians learning French than on French speakers of Italian. Moreover, the French learners would not tend to carry over their more complex native system into Italian, whereas the Italian learners would tend to carry over their less complex native system into French.

The second assumption is that Italian and French form a superset/subset relation with respect to ESSERE/ETRE selection: all verbs requiring ETRE in French also require ESSERE in Italian, but the set of ESSERE Italian verbs is broader than the set of ETRE French verbs and properly includes it.
The Subset Principle (Berwick 1986) would predict that both Italian and French learners would start from the narrowest (least marked) grammar, which is the French grammar; they would assume that the French system of ETRE assignment is the correct one. Italian learner of French would not need to revise this hypothesis, since the input would provide them with confirming evidence. French learners of Italian, on the other hand, would have to modify their starting hypothesis, which is too restrictive to accomodate for the facts of Italian: they would be able to do so on the basis of disconfirming positive evidence, which they would have no trouble in getting since the Italian system is assumed to be a superset of the French system. The Subset Principle would therefore prevent any learnability problem.

There are, however, serious doubts as to whether the Subset Principle operates at all in adult second language acquisition, because it would imply that learners proceeeed in total disregard of the facts of their native language (for extensive discussions see Flynn 1988; White 1990). It seems more plausible to think that learners - at least initially - make hypotheses that are compatible with their native language.[1]

It is precisely the assumption that the Subset Principle does *not* underlie adult language acquisition that leads us to hypothesize directional difficulties, at least in those cases in which the L1 and the L2 fall into a proper subset/superset relation. In fact, the the Subset Principle was proposed for first language acquisition as an explanation for the child's apparent ability to retreat from incorrect generalizations on the basis of positive evidence only. If the child starts from an incorrect hypothesis

that accounts for the largest grammar, there will be no positive evidence in the language being learned indicating that she is on the wrong track, since all the positive data will be generated by the (over-general) grammar. However, the order of hypotheses that the child can entertain is supposed to be constrained by the Subset Principle in such a way that the broader grammar cannot be assumed to start with. In contrast, the adult learner, lacking the Subset Principle, will face a problem in learning an L2 that is a proper subset of her L1; in contrast, the learner of a language which is a proper superset of her L1 will - in principle - encounter no problem. In the first case, persistent (and possibly permanent) errors are likely to occur.

Under the assumption that second language learners do not have access to the Subset Principle, French learners of Italian can be predicted to be in a more favourable position than Italian learners of French. This was precisely one of our working hypotheses:

(a) Italians experience greater difficulty in re-setting the parameter for ETRE-assignment in French L2 than French learners in re-setting the same parameter for ESSERE-assignment in Italian L2.
(b) difficulty is reflected in learners' inability to discriminate between the correct and the incorrect auxiliary with given verb types, which we take to be a measure of the indeterminacy of these verbs in their linguistic intuitions.
(c) the acceptability hierarchy for ESSERE-assignment hypothesized in 3.10.4 is reflected by the following:

1. where Italian is different from French:
   (i) Italian learners will show persistent indeterminacy for French equivalents of verbs which are at the core of the Unaccusative Hierarchy (i.e. change-of-state unaccusatives);
   (ii) Italian learners will show less indeterminacy for French equivalents of verbs which are peripheral along the Italian acceptability hierarchy (i.e. [+unergative alternant] unaccusatives) than for core verbs;
   (iii) French learners will first reach determinacy for those verb categories which are more determinate in Italian; their development will follow the unaccusative hierarchy.

2. where Italian is the same as French: no particular difficulty in either direction.

In order to investigate whether these differences exist, and whether they have any long-term effects, it was decided to test the linguistic intuitions of learners of both Italian and French at three different levels of proficiency.

5.1.1 Verb categories

The judgment test used in the experiment included 10 verb categories (the auxiliaries ESSERE/ETRE will be referred to as E, and AVERE/AVOIR as A):

(a) motional unergative verbs, which require A in both languages;
(b) non-motional unergative verbs, which require A in both languages;
(c) unergative verbs with unaccusative alternants, which require A in both languages;
(d) change-of-state unaccusatives, which require E in both languages;
(e) change-of-state unaccusatives which require E in Italian and A in French;[2]
(f) continuation-of-state unaccusatives, which require E in Italian and A in French (with the exception of *restare/rester*);
(g) existence of state unaccusatives, which require E in Italian and A in French;
(h) unaccusatives with transitive alternants, which require E in Italian and A in French;
(i) unaccusatives with unergative alternants, which require E in Italian and A in French;
(j) Raising verbs, which require either E or A in Italian and A in French.

Each category was represented by 3 verbs, except the category of continuation-of-state unaccusatives, which included only two verbs. Thus, there were a total of 29 verbs.
The complete list of verbs divided by category is given in Table 5.1:

TABLE 5.1: Pilot 1. Verb categories used in the judgment test (I = F: Italian and French select the same auxiliary; I ≠ F: Italian and French select different auxiliaries)

	Italian		French	
	essere	*avere*	*être*	*avoir*
unergative/motional		nuotare camminare viaggiare		nager marcher voyager
unergative/non-motional		dormire giocare lavorare		dormir jouer travailler
unergative with unaccusative alternant		rotolare correre saltare		rouler courir sauter
unaccusative/change-of-state (I = F)	cadere diventare rientrare		tomber devenir rentrer	
unaccusative/change-of-state (I ≠ F)	sparire invecchiare arrossire			disparaître vieillir rougir
unaccusative/continuation of state	restare sopravvivere		rester	 survivre
unaccusative/existence-of-state	esistere sembrare appartenere			exister sembler appartenir
unaccusative with transitive alternant	aumentare affondare cambiare			augmenter couler changer
unaccusative with unergative alternant	rotolare correre saltare			rouler courir sauter
Raising constructions	sembrare	 dovere potere		sembler devoir pouvoir

5.1.2 Subjects and design

38 subjects took part in the experiment: 20 native French learners of Italian and 19 native Italian learners of French. Subjects were at three proficiency levels: beginner (up to 1 year of study of the language), intermediate (up to 3 years), advanced (more than 4 years). They ranged in age from 19-35.

The test was administered to all subjects in Rome. French learners of Italian were tested in four different private language schools, whereas Italian learners of French were tested at the Centre Culturel Français. All the institutions that agreed to participate in this study held classes at five proficiency levels. Of the subjects who were tested, beginners were attending classes at the first level, intermediates were

at the third level, and advanced were at the fifth level. All learners at the intermediate and advanced levels had passed proficiency tests required by the schools in order to reach that level. Although we did not administer a pre-test, we therefore had sufficient evidence that subjects were actually different in proficiency, and that French and Italian subjects were different in a comparable way.

Given the exploratory (and exclusively theoretical) nature of this study, it was decided not to apply new elicitation methods. The instrument used was a written rating judgment test on a 10-point scale, ranging from 1 = completely acceptable, to 10 = completely unacceptable. The test also required subjects to express the degree of confidence in their judgments on a binary scale ('?' = uncertain; '!' = certain). This information, however, was not provided consistently by all subjects and therefore had to be discarded.

The test included 58 items, of which 29 grammatical sentences, containing the verbs listed above, and 29 sentences obtained by changing the auxiliary (AVERE/AVOIR instead of ESSERE/ETRE and vice versa). The instructions and the test sentences are reported in Appendix A.

All subjects took the test in a quiet room on the school premises. They were given a booklet that had written instructions in their native language on the front page, followed by the test sentences. No time limits were imposed, although most subjects completed the test in about 15 minutes.

The following steps were taken in the analysis of the data:

(a) in order to obtain a measure of the subjects' sensitivity to the the auxiliary requirements of different verb types, it was decided to calculate the differences between the acceptability ratings given to the correct auxiliaries and the ratings given to incorrect auxiliaries;
(b) a three-way, by-subjects, repeated-measures ANOVA (native language x proficiency level x semantic category) was run on the mean auxiliary differences;
(c) post-hoc Tukey tests of significance were applied on pairs of mean auxiliary differences. The critical value of q for all comparisons was 46.96.

5.1.3 Results

The mean auxiliary differences for the ten verb categories are reported in Table 5.2. Positive numbers indicate a preference for the correct auxiliary , negative numbers a preference for the incorrect auxiliary. The larger the size of the difference, the stronger the preference.

TABLE 5.2. Pilot 1: mean auxiliary differences (beg = beginners; int = intermediate; adv = advanced; I = Italian; F = French).

	FRENCH LEARNERS			ITALIAN LEARNERS		
	beg	int	adv	beg	int	adv
UNERGATIVE, MOTIONAL	6.48	2.25	7.00	0.89	7.11	7.56
UNERGATIVE, NON-MOTIONAL	8.33	7.38	8.67	7.11	7.89	8.72
UNERGATIVE, +UNACC. ALTERNANT	2.43	3.92	1.66	1.89	4.00	6.50
UNACCUSATIVE, I = F	7.86	7.29	8.74	3.61	4.83	7.33
UNACCUSATIVE, I ≠ F	-1.67	0.04	8.20	-4.00	-0.50	1.33
CONTINUATION-OF-STATE	-1.07	1.00	6.50	1.17	1.67	2.86
EXISTENCE-OF-STATE	-2.76	-0.88	5.20	-2.72	-1.95	4.50
+TRANSITIVE ALTERNANT	-0.95	-1.00	5.00	-3.22	-0.89	4.72
+UNERGATIVE ALTERNANT	-2.19	-1.77	1.83	0.50	0.28	0.56
RAISING	-3.57	-5.62	3.80	0.75	3.58	6.17

A graphic representation of these results is given in Fig. 5.1a ,b, displaying the histograms for Italian learners of French and French learners of Italian, respectively.

FIGURE 5.1a: Pilot 1. French learners of Italian, auxiliary preferences with intransitive verbs (I = F: Italian and French select the same auxiliary; I ≠ F: Italian and French select different auxiliaries; BEG = beginners; INT = intermediate; ADV = advanced)

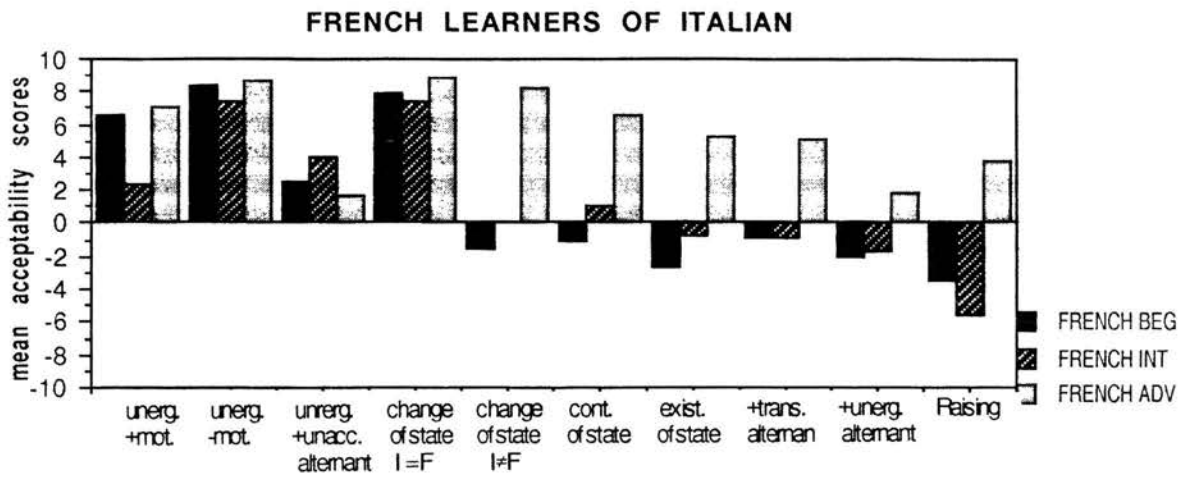
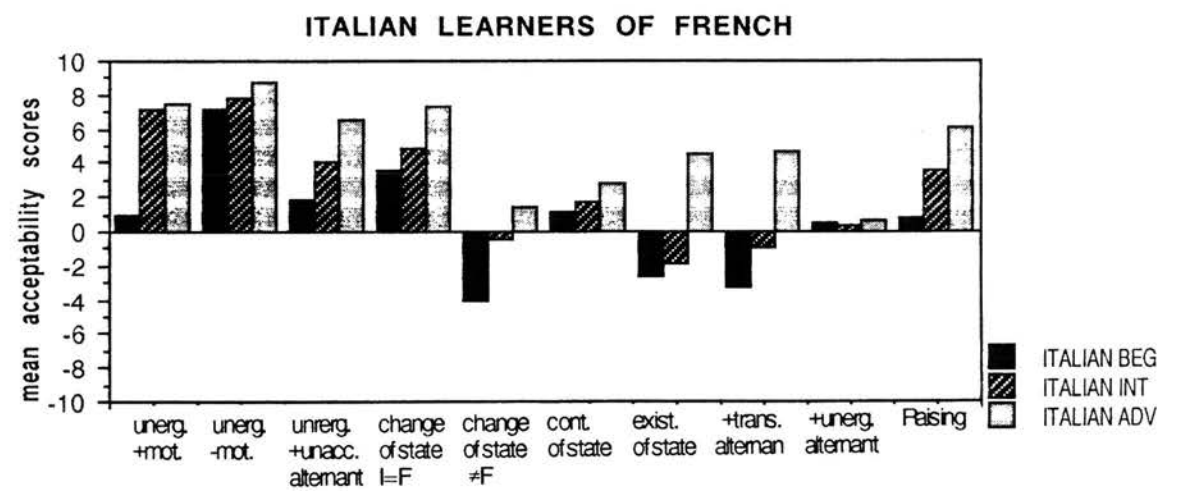


FIGURE 5.1b: Pilot 1. Italian learners of French, auxiliary preferences with intransitive verbs (I = F: Italian and French select the same auxiliary; I ≠ F: Italian and French select different auxiliaries; BEG = beginners; INT = intermediate; ADV = advanced)



The graphs in Figures 5.1a, b clearly show that judgments on different verb categories are not uniform in either language group. Statistically, this is supported in the ANOVA by a significant main effect of semantic category, $F(9, 288) = 4.30$, $p < .0001$. It is also obvious that the overall pattern of judgments is not the same across proficiency levels. The ANOVA in fact gives a significant main effect of level, $F(2, 32) = 19.62$, $p < .0001$. More interestingly, however, different verb categories elicit different responses depending on native language, proficiency level, and auxiliary.

This is reflected in the ANOVA by significant interactions of semantic category and native language, $F(9, 288) = 5.17$, $p < .0001$, semantic category and proficiency level, $F(18, 288) = 1.87$, $p < .02$, and semantic category and auxiliary, $F(9, 288) = 27.06$, $p < .0001$. The ANOVA also gives a three-way interaction of semantic category, native language and proficiency level, $F(18, 288) = 1.91$, $p < .01$, and a four-way interaction of semantic category, native language, auxiliary and proficiency level, $F(18, 288) = 1.91$, $p < .01$.

The judgments of the two language groups on unergative verb types (the three leftmost categories in Fig. 5.1a, b) are similar: in both cases, non-motional unergatives are associated with the largest auxiliary difference, followed by motional unergatives; paired unergatives (i.e. those with an unaccusative alternant) register the smallest auxiliary difference, which means that subjects do not discriminate between auxiliaries with this verbs to the same extent as with the other two verb types. Moreover, subjects appear to differentiate in similar ways between auxiliaries for change-of-state unaccusatives that select the same auxiliary (ESSERE/ETRE) in both languages. The prediction is therefore borne out that neither group would experience difficulty with respect to verb categories that require the same auxiliary in Italian and French.

Where the two groups differ most strikingly , however, is in their judgments on unaccusative verbs which select different auxiliaries in the two languages. The most obvious illustration of this are change-of-state unaccusatives. French learners of Italian have a marginal preference for the wrong auxiliary AVERE at the beginning level, are completely indeterminate at the intermediate level, and have a marked preference for the correct auxiliary ESSERE at the advanced level. Italian learners of French, on the other hand, also prefer the wrong auxiliary ETRE at the lowest proficiency level, but they remain indeterminate in their judgments through the intermediate up to the advanced level. Furthermore, the judgments of advanced

French learners of Italian are progressively less discriminating for the other types of unaccusative verbs, as indicated by the gradually decreasing sizes of the corresponding auxiliary differences. The judgments of Italian learners of French follow the reverse trend: they become progressively more determinate (as shown by the gradually increasing sizes of auxiliary differences), except for [+unergative alternant] unaccusatives. This pattern suggests that, while both groups start by assuming that the foreign language makes the same auxiliary choices as their native language, only the French learners of Italian revise their incorrect assumption, where necessary , whereas the Italian learners of French have more difficulty in recovering from initial mistakes. The degree of difficulty, for the Italian learners, appears to be proportional to the 'coreness' of verb categories along the unaccusative hierarchy.

For the Italian advanced learners (represented by the light grey bars in Fig. 5.1b), Tukey tests of significance between means give a significantly smaller auxiliary preference for change-of-state unaccusatives where I ≠ F than for non-motional unergatives (q = 48.39). Advanced learners also discriminate the two auxiliaries significantly less with [+unergative alternant] unaccusatives than with non-motional unergatives (q = 49.69) and change-of-state unaccusatives where I = F (q = 47.29). For French advanced learners (represented by the light grey bars in Fig. 5.1a) , the only significant difference is between non-motional unergatives and [+unergative alternant] unaccusatives (q = 46.28).

At the intermediate level, Italian learners (represented by the dark grey bars in Fig. 5.1b) have weaker auxiliary preferences with change-of-state unaccusatives (I ≠ F) than with non-motional unergatives (q = 46.57) and motional unergatives (q = 51.36). Auxiliary differences are also significantly smaller for existence-of-state unaccusatives than for motional unergatives (q = 55.41), and non-motional unergatives (q = 60.18). Auxiliary preferences for +transitive alternant unaccusatives are weaker than for motional (q = 48.95) and non-motional unergatives (q = 53.72). In the judgments of French intermediate learners of Italian (represented by the dark grey bars in Fig. 5.1a), auxiliary differences are significantly larger for non-motional unergatives than for existence-of-state unaccusatives (q = 50.5), +transitive alternant unaccusatives (q = 51.28), [+unergative alternant] unaccusatives (q = 55.62), and Raising verbs (q = 79.57). These results indicate that verb types that select the same auxiliary in Italian and French are associated with stronger auxiliary preferences in the judgments of both subject groups.

At the beginning level, Italian learners of French (represented by the black bars in Fig. 5.1b) have a preference for the wrong auxiliary in all verb categories that select a different auxiliary in the two languages. Thus, there are significant differences between non-motional unergatives, motional unergatives, and change-of-state unaccusatives (I = F), on the one hand, and change-of-state unaccusatives (I ≠ F) (q = 61.19; q = 49.84; q = 58.29, respectively), continuation of state unaccusatives (q = 57.55; q = 46.2; q = 54.65), existence-of-state unaccusatives (q = 67.88; q = 56.53; q = 64.98), +transitive alternant unaccusatives (q = 56.82; q = 45.47; q = 53.92), [+unergative alternant] unaccusatives (q = 64.39; q = 53.04; q = 61.49), and Raising verbs (q = 72.85; q = 61.5; q = 69.95).

French learners of Italian (represented by the black bars in Fig. 5.1a) also prefer the wrong auxiliary, but for fewer verb categories. In their judgments, there are significant differences between the auxiliary preferences with non-motional unergatives, on the one hand, and those with change-of-state (I ≠ F) unaccusatives (q = 68.01), existence-of-state unaccusatives (q = 60.19), and +transitive alternant unaccusatives (q = 63.24) on the other.

5.1.3.1 Summary of results

The findings of Pilot 1 can be summarized as follows:

Italian L1-->French L2

(a) for verb categories that select different auxiliaries in the two languages, there is little overall progress beyond intermediate level in terms of determinacy of judgments; Italian learners assume that French behaves like Italian with respect to auxiliary choice, and are reluctant to abandon this hypothesis;
(b) the Italian learners' difficulty with French verbs follows the unaccusative hierarchy, but in reverse: verb categories closer to the core of the unaccusative hierarchy are the most difficult to acquire in French L2.
(c) verbs that select the same auxiliary in the two language do not pose any particular problem.

French L1-->Italian L2

(a) there is development across levels, regardless of whether a verb category selects the same auxiliary in the two languages; French learners start by assuming that Italian behaves like French, but they correct themselves at a later stage.

(b) the variation in the judgments of French advanced learners on auxiliary selection appears to be related to the position of verb categories along the unaccusative hierarchy .

Thus, these findings reveal an asymmetrical pattern in the acquisition of auxiliary choice with unaccusative verbs. Despite the fact that both groups of learners had been exposed to metalinguistic and naturalistic input, only the French learners of Italian show evidence of having internalized the evidence, and are able to discriminate among grammatical, ungrammatical and variable auxiliary selection. The Italian learners of French, on the other hand, have difficulty in recognizing ungrammaticality in French auxiliary selection, particularly if ungrammaticality coincides with Italian auxiliary choice.

Why is there a clear directionality of difficulty for the acquisition of ESSERE/ETRE assignment , and why are French learners of Italian at an advantage over Italian learners of French?

As we have seen, two different theoretical approaches would predict these results: one considers issues of learnability such as the unavailability of the Subset Principle to second language acquirers; the other considers the semantic bases underlying the process of historical change in Romance auxiliaries, and the different effects of such process in Italian and French. We want to argue that the latter approach is the correct one, whereas the former is seriously flawed.

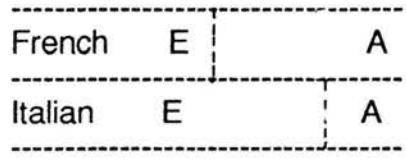
The most important issue is the relation between Italian and French with respect to auxiliary assignment. While it may seem relatively straightforward to place the two languages in a subset relation to each other, careful consideration has to be given to certain conditions that have to obtain before such a relation can be assumed. It is useful to refer specifically to what Wexler and Manzini (1987) define as the *Subset Condition* and the *Independence Principle*.

If the Subset Principle is a method for specifying a markedness hierarchy when alternative values of a parameter yield languages which are in a subset relation, it is

necessary that the two values of a parameter in fact define languages that fall into a subset relation to each other. This requirement they call the "Subset Condition". Furthermore, one needs to take into account the fact that there is more than one parameter in a language, so that just setting one parameter does not allow the language to be calculated. One has to make sure that , in the words of Wexler and Marzini, "the particular subset relation of the two languages formed by two values (say *i* and j) of a parameter are not affected by the setting of the other parameters. If *i* produces a subset of *j* for some setting of the other parameter, then it will produce the same subset for all the other values of a parameter" (p. 46). This property is called "Independence".

Now, a closer look at the facts of French and Italian disallows the possibility of the two languages satisfying either the Subset Condition or the Independence Principle. With respect to ESSERE/ETRE it is true that all verbs taking ETRE in French also take ESSERE in Italian, but with respect to AVERE/AVOIR the inverse relationship obtains. In the first case, French is a subset of Italian, whereas in the second case, Italian is a subset of French.

The relation between the two languages can be represented as in the diagram below:



This seems to be the situation envisaged by Hyams (1986) when she argues: "An interesting question raised by the Subset Principle is what happens in the case in which a value *x* along parameter P generates a 'smaller' language with respect to a range of data A, but a larger language with respect to a range of data B.....This seems to suggests that the Subset Principle has to be relativized to particular data." (Hyams 1986). In the case of the present data, the Subset Principle engenders two directly opposing predictions simply because the Subset Condition is not satisfied.

If the Subset condition is not a possible explanation for the directionality of difficulty, how can these facts be accounted for?

These findings can be explained by comparing Italian and French on a different basis. The disrupted link between unaccusativity and auxiliary selection is the most visible manifestation of a state of disequilibrium in the French auxiliary system, and is due to the fact that French has been undergoing a complex process of change. The starting point for French learners in the acquisition of Italian is a relatively inconsistent syntactic/semantic system, whereas their target is a system in which the semantic and syntactic properties of unaccusativity are still closely integrated. French learners do not experience any particular difficulty vis-a-vis Italian auxiliaries because Italian implements the 'canonical' representation of unaccusativity.[3]

On the other hand, Italian learners of French start from a canonically consistent system and have to cope with inconsistencies in the target. Thus, they experience difficulty and have indeterminate intuitions, not so much because French auxiliaries are different from their Italian counterparts, but mainly because AVOIR with unaccusative verbs (particularly with the most prototypical verbs of change of location) is unexpected and structurally inconsistent.

5.1.4 Conclusions

More generally, this study suggests the following conclusions. First, the notion of unaccusativity seems to be reflected in learners' judgments. Language learners, irrespective of the extent to which their native language conforms to the canonical implementation of unaccusativity, tend - at least temporarily - to override the second language input, if the input is inconsistent with the canonical implementation. Second, there is a hierarchy of unaccusativity that determines the relative markedness and prototypicality of different unaccusative verbs. This is reflected by the different degrees of determinacy of learners' acceptability judgments on auxiliary choice: the closer a verb category is to the core, the more determinate (i.e. consistent) acceptability judgments are; conversely, the closer a verb category is to the periphery, the less determinate judgments are.[4]

5.2 Pilot 2

Auxiliary selection and past participle agreement with transitive and intransitive verbs: three methods for the elicitation of acceptability judgments.

The second pilot study, like Pilot 1, had its main theoretical focus on the choice of auxiliary verb (AVERE / ESSERE) with different types of intransitive verbs. These included the unaccusative and unergative categories tested in Pilot 1 and, in addition, a set of 'weather verbs' that, as was noted in Chapter 3, admit an undifferentiated use of auxiliaries. Pilot 2 was concerned only with the selection of perfective auxiliaries in basic declarative sentences: it did not test intuitions on any of the syntactic correlates of the Unaccusative Hypothesis and auxiliary choice (i.e. Ne-cliticization, restructuring rules).

The secondary theoretical focus of this study was on the agreement or non-agreement of the past participle with the object NP with different types of transitive verbs. As we saw earlier (Chapter 3, section 3.8.3.2), the sub-standard acceptability of past participle agreement seems to be related to whether the verb has a resultative meaning or not: it is more acceptable with transitive verbs that denote a change of state of the direct object than with transitive verbs that do not.

Unlike Pilot 1, Pilot 2 was also concerned with methods of collecting acceptability judgments from native and non-native speakers. For this purpose , it set up a comparison between magnitude estimation and two more common procedures making use of relative judgments: card-sorting and ranking.

The theoretical background outlined in Chapter 3 again constituted the basis for the experimental hypotheses and the choice of test materials. The central idea is, once again, that there are certain clear categories where acceptability is not in dispute but that there are also unclear categories where acceptability varies among native speakers. The system of verb categories employed in Pilot 2 (like the one in Pilot 1) relies on the assumption that the classes of unergative and unaccusative verbs are not homogeneous, and that they are subdivided by lexical-semantic distinctions which can be put in hierarchical order. The corollary of this position is that core verbs require a given auxiliary in a categorical way; the further a verb category is from the core, the less categorically it selects its auxiliary.

5.2.1 Verb categories

The verb types used in the pilot study are given in Table 5.3:

TABLE 5.3: Verb categories used in Pilot 2

Category 1	Unaccusative verbs denoting change of location (*andare, uscire, arrivare*)
Category 2:	Unaccusative verbs denoting continuation of a state (*rimanere, restare, sopravvivere*)
Category 3:	Unaccusative verbs denoting existence of a state (*appartenere, piacere, sembrare*)
Category 4:	Unergative verbs denoting a locomotional activity which is not perceived as resulting in change of location (*passeggiare, nuotare, viaggiare*)
Category 5:	Unergative verbs denoting a non-locomotional activity which is not perceived as resulting in change of location (*dormire, giocare, lavorare*)
Category 6:	Intransitive verbs denoting a locomotional activity which (a) may or (b) may not be perceived as resulting in a change of state (*correre, rotolare, strisciare*)
Category 7:	Transitive verbs denoting the causation of a change of state in the object argument (*lavare, rompere, mangiare*)
Category 8:	Transitive verbs not denoting the causation of a change of state in the object argument (*baciare, suonare, stringere*)
Category 9:	Verbs which may be (a) unaccusative or (b) transitive, respectively denoting a change of state in the subject argument or the causation of a change of state in the object argument (*migliorare, aumentare, cambiare*)
Category 10:	Intransitive verbs of various semantic types which appear to take AVERE or ESSERE in free variation (*vivere, piovere, decollare*)

Although the total number of verb categories was 10, two of them in fact included two subclasses (categories 6 and 9), so that the total number of sentences included in the judgment test was 12.

This study addressed the following questions:

(a) whether there exist acceptability hierarchies in the intuitions of native Italian speakers within the classes of unergative, unaccusative, and transitive verbs;
(b) whether non-native acceptability hierarchies gradually approximate to native hierarchies and eventually - at very advanced stages of mastery of the second language - come to coincide with them;
(c) whether the three ranking procedures employed (Card-Sorting, Magnitude Estimation and straightforward Ranking) produce similar acceptability hierarchies, and whether they are equally sensitive to variations in acceptability.

5.2.2 Hypotheses

As in Pilot 1, we regard the informants' ability to discriminate between auxiliaries ESSERE and AVERE in their judgments as a measure of the determinacy of a given verb type in their linguistic intuitions.

The theoretical hypotheses, based on the acceptability hierarchies discussed in Chapter 3, predicted that:

(a) overall, judgments on past participle agreement with transitive verbs would be more determinate than judgments on auxiliary selection with intransitive verbs because past participle agreement with direct objects is a phenomenon of relative marginality;

(b) within each of the classes of unergative,unaccusative and transitive verbs, unpaired verbs (i.e. the ones without lexicalized alternants belonging to different verb classes) would be more determinate than paired verbs (i.e. the ones with such alternants).

(c) among unpaired unaccusatives, there would be a significant difference among change-of-location, continuation of state and existence-of-state verbs , with change-of-location verbs being the most determinate;

(d) among unergative verbs, there would be significant differences among motional, non-motional, and [+unaccusative alternant] unergative verbs: non-motional verbs would be most determinate, whereas [+unaccusative

alternant] verbs would be perceived as more closely associated with unaccusatives, and therefore as least determinate.

(e) there would be a significant difference among (a) transitive verbs not implying a change of state in the object, (b) transitive verbs implying a change of state in the object, and (c) paired transitive verbs with an unaccusative alternant: sub-standard past participle agreement would be most acceptable in (c) and least acceptable in (a).

(f) 'weather verbs' (*vivere, piovere, decollare*) would be in free variation (and therefore totally indeterminate) in native Italian intuitions.

(g) the acquisition of auxiliary selection by English learners of Italian (as indicated by the change in their linguistic intuitions across proficiency levels) would be easier for more determinate verb types and more difficult for less determinate verb types.

5.2.3 Subjects

Four groups of subjects participated in this study: (a) Italian native speakers; (b) near-native speakers of Italian; (c) advanced learners; (d) beginning learners.
The native speakers were all monolingual Italians living in Edinburgh, either permanently or temporarily, in the age range 19-52. The near-native speakers were interviewed individually prior the experimental sessions, and selected on the criterion that they could pass for natives with respect to fluency, accuracy and lexical choice (phonological accuracy was not taken into account); none was of Italian origin; all lived and worked in Edinburgh.
Learners were students of Italian at the Institute for Applied Language Studies of the University of Edinburgh. All non-native speakers had English as their mother tongue. None had Italian origins.
All subjects participated in the experiment on a voluntary basis.

The number of subjects at each level was as follows:

- 18 natives
- 5 near-natives

- 6 advanced learners (with more than three years of exposure to Italian, of which one spent in Italy)
- 6 low interemediate learners (with less than one year of exposure to Italian)

Subjects had to judge 72 sentences, divided into three sets, each set consisting of 24 sentences: 12 grammatical sentences, representing the verb categories identified for auxiliary choice and past-participle agreement, and 12 lexically identical counterparts, in which intransitive verbs appeared with ESSERE Instead of AVERE and vice versa, or in which transitive verbs took past participle agreement. These latter 12 sentences had varying degrees of unacceptability, depending on the verb type. The three sets were structurally equivalent to each other, since they contained different tokens of the same sentence types, as shown in the examples in (a):

(a) **Unaccusative verbs denoting change of location**

Lex 1 Maria è andata in ufficio a piedi
Lex 2 Paola è uscita di casa alle nove
Lex 3 I miei genitori sono arivati in macchina

The three sets of sentences will henceforth referred to as Lex1, Lex 2, and Lex 3. Sentences were presented in a different random order within each lexicalization. The test sentences can be found in Appendix A.

5.2.4 Tasks

Subjects were asked to perform in three different types of judgment test:

(a) **Ranking**: subjects were presented with a list of 24 sentences in random order and required to rank them according to their degree of acceptability, starting from 1 = most acceptable. Ties were allowed. No time limits were imposed.

(b) **Magnitude Estimation**: Subjects were presented with 24 isolated sentences, projected on a screen one at a time by means of an overhead projector. They had an answer sheet in front of them. They were instructed to give the first sentence any number they wished and then to assign numbers to successive sentences in a way that reflected their acceptability relative to the first sentence.

(c) Card-Sorting: while sitting at a desk, subjects were given a set of 24 cards, each having a numbered sentence written on it. They also had a tape measure in front of them, stretched out along the desk. Subjects were instructed to sort the cards out into piles according to their degree of acceptability, ranging from '-acceptable" on their left to "+acceptable" on their right, and then to reproduce the order and content of the piles along the tape measure, by transcribing the numbers corresponding to sentences in each pile in separate columns. They were free to form as many piles as they wished and to place the piles at any position they thought appropriate. Again, no time limits were imposed.

The complete set of test materials can be found in Appendix A.

Because of unexpected difficulties in arranging suitable times for the experiment, advanced and low intermediate learners were only administered the Ranking and Magnitude Estimation tests. Near-native speakers performed in all three tasks as planned.

Subjects were tested on university premises. They sat in a quiet room, equipped with an overhead projector and desk space. Written instructions (in Italian for native speakers and in English for non-native speakers) were provided at the beginning of the experiment, followed by further verbal clarification if required. The word "acceptability" was used throughout the instructions, and its interpretation was deliberately left vague in order to avoid an exclusive focus on grammatical correctness. No time limits were imposed. Subjects generally took approximately 30 minutes to complete the battery of tests.

At the end of the session, they were invited to express informal comments on the tests. These comments indicated a preference for Card-sorting, while Magnitude Estimation and Ranking had both a low degree of face-validity. Most subjects expressed the view that Magnitude Estimation was 'impossible' and felt that they had not been able to follow the instructions. Among the most frequent comments were the following:

- subjects could not remember the number assigned to the first sentence;
- subjects could not think in terms of ratios and proportions.

Ranking was also criticized for causing 'confusion', because all the test sentences were presented at the same time and subjects could not focus on each sentence separately.

5.2.5 Experimental design

The sample of subjects consisted of two groups of 18 native speakers and 18 non-native speakers. Both groups were divided into 3 sub-groups A, B and C of 6 people each. In the case of the non-native speakers, this division coincided with the three levels of proficiency, so that each level consisted of 6, 6, and 5 subjects respectively.

Each sub-group received a different combination of tests and lexicalizations, as can be seen in the Latin square below:

	Lex 1	Lex 2	Lex 3
Magnitude Estimation	**C**	**B**	**A**
Card-Sorting	**A**	**C**	**B**
Ranking	**B**	**A**	**C**

Each lexicalization of each verb type was judged by each method by 1/3 of each proficiency group. Each sub-group was further divided into thirds and each third received the ranking tasks in a different order. Assignment of lexicalizations to groups and tasks to orders was by Latin square. The result was a design balanced for order of presentation of stimuli and of experimental tasks.

The data were subjected to the following analyses:

(a) A s in Pilot 1, the differences were calculated, for each subject and each sentence type, between responses on correct sentences (containing the correct auxiliary/agreement) and those on incorrect sentences (containing the wrong auxiliary/agreement). Arithmetic means were then computed on the resulting "preference scores" for Ranking and Card-Sorting and geometric means for Magnitude Estimation.

(b) a series of by subject, repeated measures ANOVAs (proficiency level x verb category) were performed separately on the mean auxiliary differences resulting from Magnitude Estimation,Card-Sorting and Ranking.

(c) post-hoc Tukey tests of significance were applied to paired comparisons between mean auxiliary differences. The following comparisons were examined:

i. among verb categories within each proficiency level

ii. among proficiency levels for each category.

The critical values of q for all comparisons in each method are as follows:

Magnitude Estimation: q = 5.29

Card-Sorting: q = 17

Ranking: q = 37.44.

5.2.6 Results

The mean preference scores for all verb types are given in Table 5.4 (Magnitude Estimation), Table 5.5 (Card-Sorting), and Table 5.6 (Ranking).

TABLE 5.4. Pilot 2: mean preference scores (in logarithmic form) in Magnitude Estimation judgments on intransitive and transitive verbs (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

	INT	ADV	NNS	NS
UNERGATIVE, NON-MOTIONAL	0.519	0.786	0.876	0.879
UNERGATIVE, MOTIONAL	0.082	0.252	0.531	0.817
UNERGATIVE, +UNACC. ALTERNANT	-0.088	0.128	0.190	0.334
UNACCUSATIVE, CHANGE OF STATE	-0.159	-0.085	0.520	0.805
UNACCUSATIVE, CONT. OF STATE	0.391	0.381	0.348	0.555
UNACCUSATIVE, EXISTENCE OF STATE	0.005	0.583	0.373	0.503
UNACCUSATIVE, +TRANSITIVE ALTERNANT	-0.175	0.205	0.295	0.358
UNACCUSATIVE, +UNERGATIVE ALTERNANT	0.066	-0.204	-0.080	0.132
WEATHER VERBS	0.125	-0.380	-0.216	0.125
TRANSITIVE, CHANGE OF STATE IN DO	0.182	0.600	0.348	0.434
TRANSITIVE, NO CHANGE OF STATE IN DO	0.230	0.705	0.800	0.766
TRANSITIVE, +UNACC. ALTERNANT	-0.204	0.724	0.835	0.801

TABLE 5.5. Pilot 2: mean preference scores in Card-Sorting judgments on transitive and intransitive verbs (NNS = near-native speakers; NS = native speakers).

	NNS	NS
UNERGATIVE, NON-MOTIONAL	3.200	2.444
UNERGATIVE, MOTIONAL	2.600	2.278
UNERGATIVE, +UNACC. ALTERNANT	1.600	1.222
UNACCUSATIVE, CHANGE OF STATE	2.400	2.167
UNACCUSATIVE, CONT. OF STATE	1.800	2.111
UNACCUSATIVE, EXISTENCE OF STATE	1.800	2.056
UNACCUSATIVE, +TRANSITIVE ALTERNANT	1.800	1.500
UNACCUSATIVE, +UNERGATIVE ALTERNANT	-1.000	-0.167
WEATHER VERBS	-0.400	-0.333
TRANSITIVE, CHANGE OF STATE IN DO	1.800	2.167
TRANSITIVE, NO CHANGE OF STATE IN DO	1.800	2.944
TRANSITIVE, +UNACC. ALTERNANT	2.600	2.944

TABLE 5.6: Pilot 2.: mean preference scores in Ranking judgments on intransitive and transitive verbs (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers).

	INT	ADV	NNS	NS
UNERGATIVE, NON-MOTIONAL	3.667	6.833	3.600	5.500
UNERGATIVE, MOTIONAL	-2.167	2.500	3.000	4.222
UNERGATIVE, +UNACC. ALTERNANT	-2.667	-1.000	2.000	3.389
UNACCUSATIVE, CHANGE OF STATE	2.000	-1.000	2.200	5.222
UNACCUSATIVE, CONT. OF STATE	2.500	4.167	1.600	3.833
UNACCUSATIVE, EXISTENCE OF STATE	4.000	6.167	1.800	3.889
UNACCUSATIVE, +TRANSITIVE ALTERNANT	0.000	3.000	1.800	3.167
UNACCUSATIVE, +UNERGATIVE ALTERNANT	1.667	0.333	-2.200	0.611
WEATHER VERBS	0.000	0.667	-0.200	0.000
TRANSITIVE, CHANGE OF STATE IN DO	4.000	5.333	1.600	3.500
TRANSITIVE, NO CHANGE OF STATE IN DO	1.000	2.167	1.600	5.167
TRANSITIVE, +UNACC. ALTERNANT	4.500	6.667	2.800	5.611

Visual representation of the data is given in Fig. 5.2a, b, Fig. 5.3a, b, and Fig. 5.4a, b. In all figures, the (a) graphs represent the mean auxiliary preferences for intransitive verbs; the (b) graphs represent the mean agreement preferences for transitive verbs.

FIGURE 5.2a. Pilot 2: mean auxiliary preferences in **Magnitude Estimation** judgments on **intransitive** verbs (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

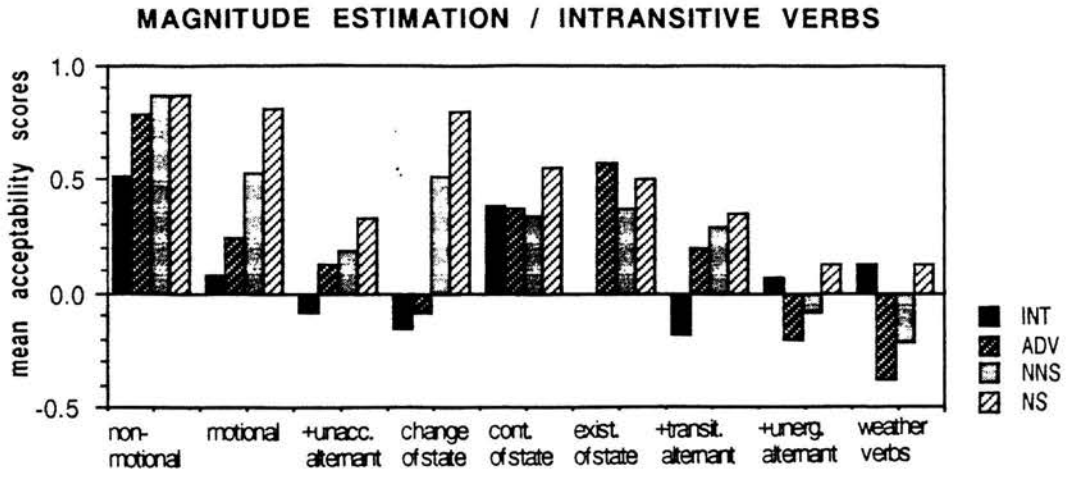


FIGURE 5.2b. Pilot 2: mean agreement preferences in **Magnitude Estimation** judgments on **transitive** verbs (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

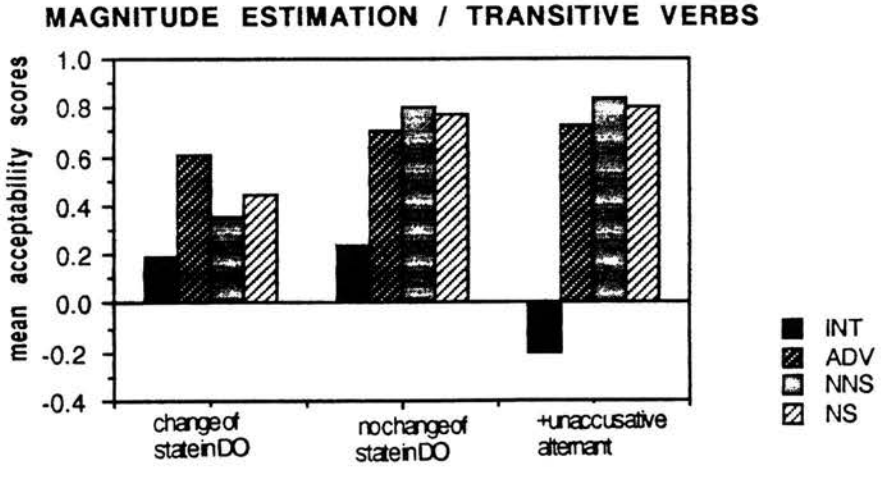


FIGURE 5.3a. Pilot 2: mean auxiliary preferences in **Card-Sorting** judgments on **intransitive** verbs (NNS = near-native speakers; NS = native speakers)

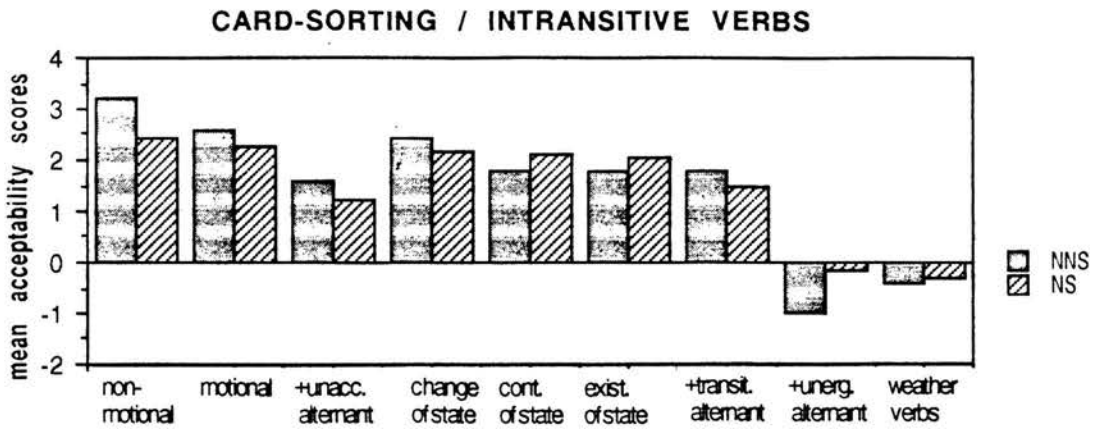


FIGURE 5.3b. Pilot 2: mean agreement preferences in **Card-Sorting** judgments on **transitive** verbs (NNS = near-native speakers; NS = native speakers)

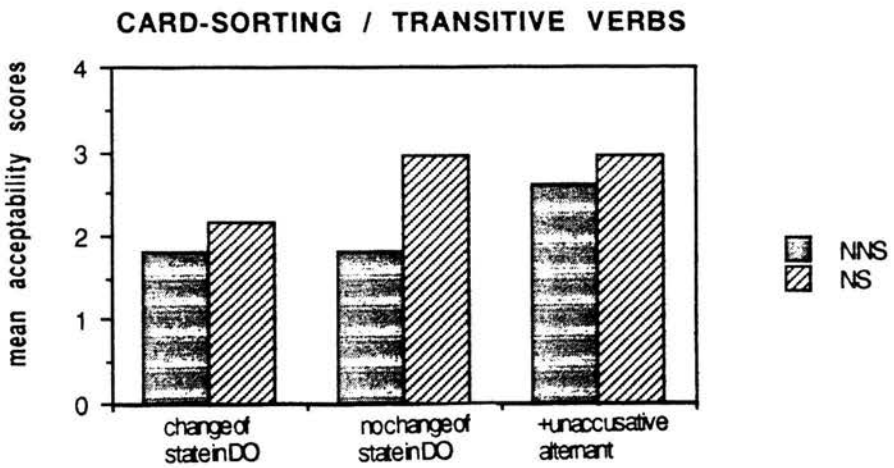


FIGURE 5.4a. Pilot 2: mean auxiliary preferences in **Ranking** judgments on **intransitive** verbs (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

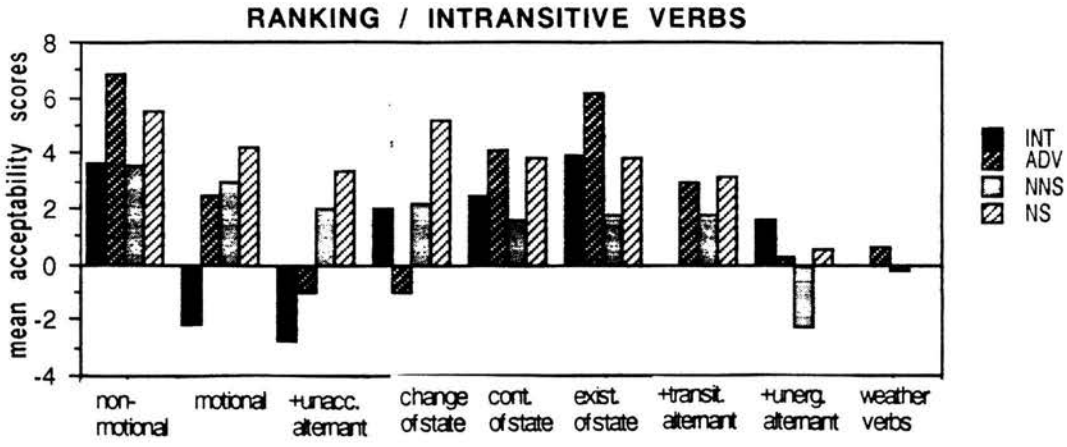
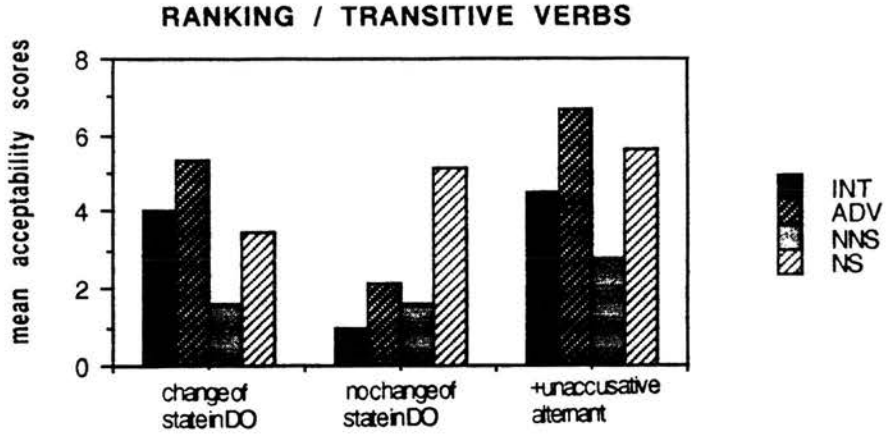


FIGURE 5.4b. Pilot 2: mean agreement preferences in **Ranking** judgments on **transitive** verbs (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)



The figures for all the three methods generally show that the informants' preference scores, and the strength of these preferences, are not uniform: they vary according to the verb categories. This is confirmed by a significant main effect of semantic category in the ANOVAs for all three methods: Magnitude Estimation,
$F(11, 341) = 12.14$, $p < .0001$, Card-Sorting, $F(11, 231) = 16.37$, $p < .0001$, Ranking, $F(11, 341) = 11.25$, $p < .0001$. For Magnitude Estimation and Ranking (Figs. 5.2, 5.4) variation is affected by proficiency level; proficiency is less relevant to the Card-Sorting data (Fig. 5.3), since the test was administered only to native and near-native speakers. Both Magnitude Estimation and Ranking show a significant interaction of proficiency level and semantic category: $F(33, 341) = 2.07$, $p < .0007$ and $F(11, 341) = 2.83$, $p < .0001$, respectively. Only Magnitude Estimation, however, registers a significant main effect of proficiency level, $F(3, 341) = 6.04$, $p < .002$. Predictably, the ANOVA on the Card-Sorting data does not give either a main effect of proficiency level, or an interaction with proficiency level, suggesting that native and near-native informants perform similarly.

5.2.6.1 Acceptability hierarchies in native judgments

The rank orders of the same sentences produced by native speakers in the three methods present broad similarities, as can be seen in Tables 5.2-5.4. These orders, for each verb class, indicate that:

(a) preference scores are generally stronger within transitive and unergative verbs than within unaccusative verbs;
(b) for unergative verbs, auxiliary preferences for AVERE are stronger with non-motional verbs than with motional verbs, and weakest with [+unaccusative alternant] verbs;
(c) for unaccusative verbs, preference scores for ESSERE (in Magnitude Estimation and Card-Sorting, but not Ranking) are strongest with change-of-state verbs and weakest with [+unergative alternant] verbs. The strength of preferences for continuation-of-state, existence-of-state and +transitive alternant is generally, but not always, predictable from their relative order on the unaccusative hierarchy proposed.
(d) for transitive verbs, preferences for agreement with the direct object are only marginally stronger with [+unaccusative alternant] verbs than with [+change-of-state-in DO] verbs and [-change-of-state-in DO] verbs. The differences

among preference scores for the three transitive verb types do not reach statistical significance.

(e) auxiliary choice with 'weather' verbs is the most indeterminate.

These similarities with respect to the general rank ordering of verb categories, and their general conformity to the expected orders suggest that the three methods produce roughly similar results. However, it can be easily seen in Fig. 5.3 and 5.4 that the Card-Sorting and Ranking graphs have a 'flatter' profile than the Magnitude Estimation graphs. A quick glance at the figures in Tables 5.3 and 5.4 (representing Card-Sorting and Ranking) indicates that several categories have the same mean preference scores, whereas this is not the case in Table 5.2 (representing the Magnitude Estimation data). The methods clearly differ with respect to their sensitivity to degrees of acceptability, as the numbers of comparisons significant by the Tukey tests show.[5]

The Magnitude Estimation data give a significantly weaker preference scores with 'weather' verbs than with non-motional unergatives (q = 5.8), motional unergatives (q = 5.38), change-of-state unaccusatives (q = 5.3), [+unaccusative alternant] transitives (q = 5.61), and no=change-of-state-in-DO transitives (q = 5.37). Similarly, preference scores with [+unergative alternant] unaccusatives are significantly weaker than those with non-motional unergatives (q 5.39), motional unergatives (q = 5.33), change-of-state unaccusatives (q = 5.28), and [+unaccusative alternant] transitives (q = 5.3).

A similar pattern is given by the Card-Sorting data. Preference scores with 'weather' verbs are significantly smaller than those with non-motional unergatives (q = 21.73), motional unergatives (q = 20.43), change of state unaccusatives (q = 19.55), and all the transitive verb types (change-of-state-in DO: q = 19.55; no-change-of-state-in-DO: q = 25.64; [+unaccusative alternant]: q = 25.64).

The Ranking data for native speakers, however, show only a significantly smaller preference score with 'weather' verbs than with non-motional unergatives (q = 37.35) and for [+unaccusative alternant] transitives (q = 38.09). No significant difference is found among the other verb categories.

5.2.6.2 Acceptability hierarchies in non-native judgments

Let us now turn to the non-native judgments. Tukey tests on the Magnitude Estimation data show that the overall number of significant differences increases with proficiency, thus indicating an increased ability to perceive degrees of acceptability. The overall number of significant differences was, however, lower than we expected.

Near-natives have a significantly more marked auxiliary preference with non-motional unergatives than with [+unergative alternant] unaccusatives (q = 6.49) and 'weather' verbs (q = 7.42). Preference scores with [+unaccusative alternant] transitives are higher (and judgments therefore more determinate) than those with [+unergative alternant] unaccusatives (q = 6.21) and 'weather' verbs (q = 7.14).
At the advanced level, there is a significantly preference score with non-motional unergatives than with [+unergative alternant] unaccusatives (q = 6.73) and 'weather' verbs (q = 7.92). No significant differences among verb categories are found at the intermediate level.

The Card-Sorting data from near-native speakers are consistent with this pattern. Preference scores are smaller with [+unergative alternant] unaccusatives than with non-motional unergatives (q = 32.84), motional unergatives (q = 28.15), and [+unaccusative alternant] transitives (q = 28.15). Preference scores with 'weather' verbs are significantly smaller than with non-motional unergatives (q = 28.15), motional unergatives (q = 23.46) , and [+unaccusative alternant] transitives (q = 23.46).

The Ranking data from non-native informants offer fewer significant differences overall, but the advanced group have generally larger preference scores than the other two proficiency levels. Furthermore, some of these differences are theoretically unexpected and inconsistent with the pattern given by Magnitude Estimation and Card-Sorting. Advanced learners in fact have significantly stronger preferences with non-motional unergatives than with change-of-state unaccusatives (q = 53.18), [+unaccusative alternant] unergatives (q = 53.18), [+unergative alternant] unaccusatives (q = 44.13), and 'weather' verbs (q = 41.86). Their preference scores are also significantly more marked for +existence-of-state unaccusatives than for change-of-state unaccusatives (q =

48.66), [+unaccusative alternant] unergatives (q = 48.66), and [+unergative alternant] unaccusatives (q = 39.61). The advanced learners' pattern of preferences contrasts sharply with those of near-native and intermediate informants. Near-natives, in fact, only have a stronger preference score with non-motional unergatives than with [+unergative alternant] unaccusatives (q = 39.38). Intermediate learners distinguish between non-motional and [+unaccusative alternant] unergative verbs (q = 43.01, and between [+unaccusative alternant] transitive verbs and [+unaccusative alternant] unergatives (q = 48.67).

Overall, these findings suggest that the results of Magnitude Estimation and Card-Sorting are more consistent with each other than with the results of Ranking.

5.2.6.3 Developmental aspects

The only comparisons relevant to development across proficiency levels are those between Magnitude Estimation and Ranking, since the Card-Sorting test was not administered to the two lowest proficiency groups.
These comparisons do not produce many significant differences. However, in the Magnitude Estimation data, Tukey tests give a significant difference between intermediate learners and native speakers in their auxiliary preferences on motional unergatives (q = 5.3), change-of-state unaccusatives (q = 6.55), and [+unaccusative alternant] transitives (q = 6.82). None of these differences, or any other differences, are found in the Ranking data, suggesting once again that Magnitude Estimation is the more sensitive method of the two.

5.2.7 Conclusions

The overall pattern of results suggests that Magnitude Estimation may be an adequate instrument for the collection and analysis of acceptability data. Pilot 2 showed that this method, despite its lower face-validity, is at least as revealing as an untimed ranking procedure such as Card-Sorting, often more sensitive to subtle acceptability distinctions, and more likely to capture variation in acceptability judgments. Although the results of Pilot 2 do not provide very strong evidence for the existence of the predicted acceptability hierarchies, or for the gradual approximation of non-native to native hierarchies, the results are encouraging because Magnitude Estimation:

(a) reveals effects that make sense from a linguistic point of view: it confirms the prediction that the classes of unergative and unaccusative verbs are not homogeneous in the intuitions of native Italian speakers;
(b) shows some systematic differences between native and non-native acceptability judgments;
(c) produces results that include most of the results of the other elicitation methods.

5.3 Pilot 3

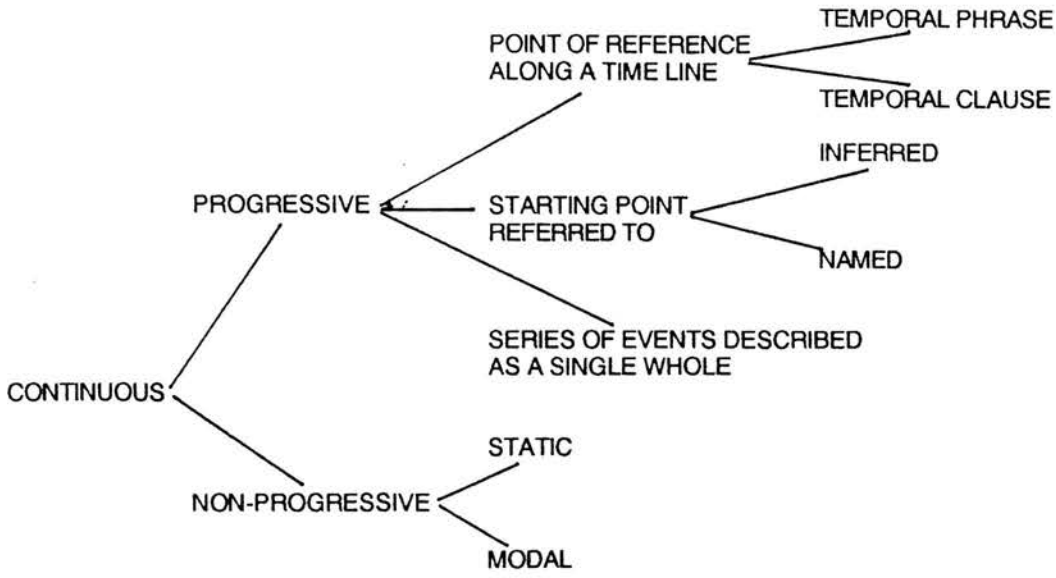
The perfective/imperfective distinction in Italian

The third pilot study addressed the same theoretical and methodological questions as Pilot 2, but was concerned with a different area of grammar: the aspectual distinction between the imperfect and the present perfect tenses in Italian. The following discussion will deal, in a necessarily concise way, with the fundamentally aspectual differences between imperfective and perfective forms. Reference will be made to a system of semantic aspectual categories which are formally expressed in Italian and a schematic comparison will be carried out with the corresponding system of categories in English.

The main semantic distinctions that will be introduced in this section are summarized by the chart in Table 5.7, where each of the end branches can be seen as representing a particular category of use within the imperfective and the perfective areas. Each category is in turn described as a parameter, or set of parameters, characterizing the situation to which the category applies.

TABLE 5.7: Pilot 3. Aspectual categories

(a) IMPERFECTIVE



HABITUAL	SITUATION CHARACTERISTIC OF WHOLE PERIOD IN THE PAST ONGOING SITUATION IN LIMITED PERIOD OF TIME ITERATIVE SITUATION (FREQUENCY/NO. OF TIMES UNSPECIFIED)

(b) PERFECTIVE

A. DURATION SPECIFIED
B. END-POINT OF SITUATION SPECIFIED
C. SITUATION WITHOUT INTERNAL STRUCTURE
D. ITERATIVE SITUATION (FREQUENCY/NO. OF TIMES SPECIFIED)
E. ITERATIVE SITUATION (TEMPORAL MARKERS REFER TO THE WHOLE)

It is our assumption that not all parameters enjoy the same status, in that - as Dahl (1985) puts it - the extension of a category has a *focus* and a *periphery*, whereby only those parameters defining the focus fully belong to the category, whereas the parameters defining the periphery have dubious membership. Whether a given parameter is basic or secondary can be either a theoretical question (i.e. it can be predicted from a particular theory), or an empirical question that can be approached on the basis of the intuitions of native speakers. (One would of course expect a positive correlation between the results of the two approaches). If the latter approach is chosen, the empirical question concerns the possibility of obtaining an acceptability hierarchy along which to place the different categories of use within the perfective/imperfective areas, such that sentences exemplifying focal parameters are judged as acceptable by native speakers with the maximum

degree of inter- and intrasubject consistency, whereas sentences exemplifying peripheral parameters are judged either as acceptable or as unacceptable with a higher degree of variability and inconsistency.

As a working hypothesis, one can speculate that the focal parameters are those defining situations which unambiguously require the use of a given form (either a perfective or an imperfective form). Conversely,peripheral parameters apply to those situations that allow for both perfective and imperfective forms. Also, one may ask as an empirical question whether incompletion and habituality, or duration and internal structure of a situation, have an equally focal status in the speaker's perception of imperfectivity or perfectivity, and whether this status is also perceived by non-native speakers in the process of language acquisition.

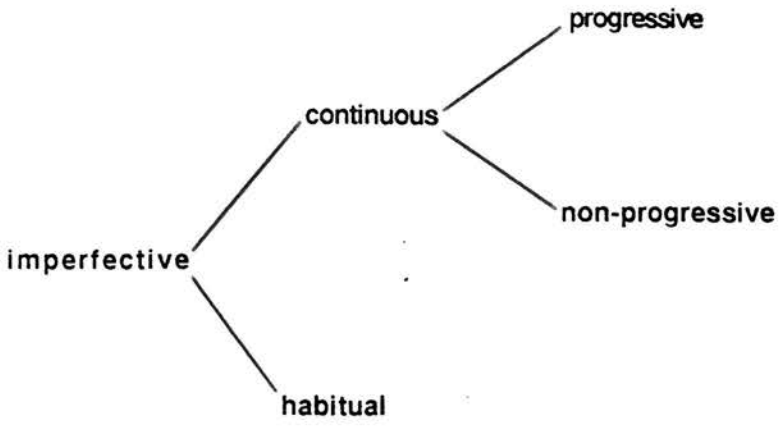
If it were indeed the case that aspectual parameters differ from one another, one would expect erroneous sentences resulting from violations to parameters to fall along a continuum, where violations to focal parameters would be considered as unacceptable by the majority of native speakers, and violations to peripheral parameters would have a fuzzier acceptability status, triggering a high degree of variability.

Let us now examine the semantic system underlying this area of grammar in some detail. This analysis will set up the framework for the experimental hypotheses that were tested in this study.[6]

5.3.1 Imperfectivity

According to Comrie (1976), the basic feature of perfectivity is "the view of a situation as a single whole, without distinctions of the various phases that make up that situation". The internal structure that a given situation may have does not enter into the meaning expressed by perfective forms. This is, on the other hand, what characterizes imperfective forms, which look at situations without considering their beginning or end, and which can indeed be associated with situations without any intrinsic temporal boundaries.

Imperfectivity can be subdivided as follows:



Habituality describes a series of iterative situations, whereas continuousness defines a single situation lasting in time and described in terms of its phasal sequence. Progressivity is typically associated with dynamic situations, as opposed to static, or non-progressive, situations. In general, Italian grammaticalizes the distinction between perfective and imperfective meaning to a much greater extent than English.

5.3.1.1 Continuous/progressive

In Italian, both the imperfect tense and the periphrastic progressive form "*stare* + gerund" can be used to express the "ongoingness" of a situation characterized as having already started but which is not yet terminated. With situations of this type, implying extension in time, a point of reference is usually specified, either in the form of a temporal phrase:

(1) Ieri a quest'ora dormivo
'Yesterday at this time I was sleeping'

or in the form of a temporal clause:

(2) Quando ha telefonato Maria, guardavo la TV
'When Maria telephoned, I was watching TV'

The situation is seen as stretching in time around these temporal markers. If temporal boundaries are specified, the imperfect cannot be used, unless the

situation is interpreted as habitual (notice the difference between (3) and (4) below):

(3) *Leggevo per tre ore
'I was reading for three hours

(4) Leggevo per tre ore ogni giorno
'Every day I read for two hours'

One can only specify the period within which the situation takes place:

(5) Sabato mattina ero nel mio ufficio
'On Saturday morning I was in my office'

The imperfect can be used when only the starting point of the situation is referred to. The starting point may be inferred:

(6) Aspettavo l'autobus da dieci minuti quando ho visto Paolo
'I had been waiting for the bus for ten minutes when I saw Paolo'

or explicitly named:

(7) Guidavo dalle sei di mattina quando ho avuto un incidente
'I had been driving since 6 'o clock in the morning when I had an accident'

In these cases what is specified is that part of the situation which has already started and which is retrospective with respect to the main happening.
The meaning of incompletion and ongoingness can also apply to a situation which includes a series of events described as a single whole:

(8) Ieri alle due distribuivo manifesti in via Cavour
'Yesterday at two I was handing out circulars in via Cavour'

The English progressive is characterized by the same notions underlying the Italian progressive. Both languages use the same temporal markers with the progressive, and in both the use of the progressive is incompatible with the use

of temporal markers specifying the duration of the action or situation. The difference lies in the expression of situations for which the starting point is specified: in English, the past perfect progressive is used. This form expresses the retrospective extension of a situation prior to its primary point of reference (I had been driving for six hours when I felt sick).

5.3.1.2 Continuous/non progressive

The imperfect is also used to express situations that are intrinsically incompatible with any temporal specification. These are typically static situations, the duration of which is not measurable, since no idea of completion or incompletion can be associated with them:

(9) La montagna sorgeva tra due valli
'The mountain was situated between two valleys'

(10) Dopo la passeggiata avevamo fame
'After the walk we were hungry'

A whole class of verbs are intrinsically imperfective, since they refer to temporally unrestricted situations:

(11) Maria ricordava tutti i particolari del suo matrimonio
'Maria remembered all the details of her wedding'

These verbs can always be transformed into perfective forms if temporal limitations are introduced:

(12) Maria ricordò quel fatto per molti anni
'Maria remembered that fact for many years'

The range of "modal" uses of the imperfect tense is interesting because it emphasizes the "non-actuality" or simply "virtuality" of the past action referred to. Consider the following sentence:

(13) *Venivo* a chiederle un favore
'I was coming to ask you a favour'

The use of the imperfect produces an effect of "vagueness" which offers the interlocutor the opportunity not to answer directly. The use of the present tense ("*Vengo a chiederle un favore*") would in fact sound more straighforward and rather rude.

The evocation of a distant or hypothetical world, somewhat detached from reality, is even more obvious in other modal uses of the imperfect, such as the "ludic" imperfect, which is often used by children to indicate the assumption of roles in play, as in the following:

(14) Facciamo che io *ero* il poliziotto e tu *eri* il ladro
'Let's pretend I was the policeman and you were the thief'

Other common modal uses of the imperfect are the potential and the hypothetical, which often replace the conditional to stress the character of inactuality or supposition of the situation. As to the former, the modal verbs *dovere, volere* and *potere* can be used in the imperfect to describe an event that was supposed to take place, but did not:

(15) Dovevo partire, ma ho perso il treno
'I was supposed to leave but missed the train'

Perfective forms would be ungrammatical in this case (**Sono dovuto partire).* Also, consider the following:

(16) Dovevo essere in ufficio tra due minuti
'I had to be in the office in two minutes'

for which conventional grammars would prescribe the past conditional (*Avrei dovuto essere..)* but the imperfect is more common in the spoken language. Finally, the imperfect can be used in contrary-to-fact conditionals both in the protasis (instead of the subjuntive), or in the apodosis (instead of the conditional), or in both:

(17) Se *studiavi*, avresti superato l'esame
(18) Se avessi studiato, *superavi* l'esame
(19) Se *studiavi*, *superavi* l'esame
'If you had studied, you would have passed the exam'

These uses of the imperfect are still banned by prescriptive grammars, although they are widespread in informal registers. No hypothetical or ludic imperfects were included in the judgment test, since learners of Italian are usually taught that these forms are unacceptable. Only examples such as (15) above were included.

English has no specific aspectual markers to express temporally unrestricted, static situations. There is therefore no difference between those static situations for which no temporal limitations are given and those for which temporal boundaries are specified (examples 12,13 above).

5.3.1.3 Habitual

The imperfect can be associated with situations that are completed, provided that these are habitual. As a general characteristic, neither the duration, nor the total number of times a situation is repeated are specified.
Within the area of habituality one can find:
(a) situations characteristic of a whole period referred to in the past, where the frequency is not specified:

(20) Da giovane Mario portava i capelli lunghi
'When he was young Mario had long hair'

(b) situations characteristic of a limited period of time, which also allow the use of forms not associated with habituality:

(21) In quel periodo Giovanna stava dai nonni
'During that period Giovanna stayed with her grandparents'

(c) iterative situations, where the repetition of the action is specified by the use of adverbial markers of frequency but the number of times or the period within which the repetition takes place are left unspecified:

(22) L'estate scorsa andavamo al mare tutti i fine settimana
'Last summer we went to the beach every weekend'

In English, habitual and iterative situations are not marked aspectually. There is therefore no difference between perfective situations for which time limits are

specified and imperfective situations for which no temporal boundaries are given ("My brother went to school by bike for ten years" and "My brother went to school by bike every day). These two situations are expressed differently in Italian, as will be seen below.

5.3.2 Perfectivity

Two tenses may be used to express perfective meaning in the past: the simple past (*passato remoto)* and the present perfect (*passato prossimo).*
Whereas the former is more used in the written language and is particularly associated with the literary language, the latter is widely used in the spoken language (especially in central and northern Italy). This study is only concerned with the present perfect.
Perfective forms generally convey a meaning of completion, that is found in situations not conceivable as having an internal structure:

(23) Il gatto ha rotto un bicchiere
'The cat broke a glass'

The perfective is also used to express situations which have a beginning, an end and a duration. The limits of the situations may be explicit:

(24) Ho avuto una casa al mare per due anni
'I had a house at the seaside for two years'

The limits are also made explicit if the end-point of the situation is given:

(25) Ieri ho dormito fino a mezzogiorno
'Yesterday I slept until noon'

The perfective area includes situations for which the frequency is determined:

(26) Stamattina ho telefonato a Mario due volte
'This morning I telephoned Mario twice'

One also finds perfective forms used in iterative situations for which the actual frequency is not specified, but the temporal limits within which the situation takes place are expressed:

(27) Per sei mesi sono andato a lezione di italiano tutti i giorni)
'For six months I went to Italian classes every day'.

The same range of situations which are expressed by perfective forms in Italian are expressed by the simple past in English. As mentioned before, English does not distinguish between imperfective situations without any temporal restrictions and perfective situations with temporal boundaries. So, for example, (28) and (29) are both expressed by the past tense, whereas Italian marks imperfectivity by using the imperfect tense:

(28) I had a house in the mountains
'Avevo una casa in montagna'

(29) I had a house in the mountains for two years
'Ho avuto una casa in montagna per due anni'

The same considerations apply to iterative situations for which either the frequency or the temporal limits are specified and situations where the repetition of the single event, or the period within which it takes place are not determined: they are distinguished in Italian but not in English:

(30) Maria went to see Paolo every day for two years
'Maria è andata a trovare Paolo ogni giorno per due anni'

(31) Maria went to see Paolo three times last week
'Maria è andata a trovare Paolo tre volte la settimana scorsa'

(32) Every day Maria went to see Paolo
'Ogni giorno Maria andava a trovare Paolo'

5.3.3 Hypotheses

Our hypotheses were centred on the idea that the distinction incompletion vs completion of an action is the most salient aspectual feature that separates imperfectivity from perfectivity. It was therefore predicted that:

(a) judgments on sentences including the imperfect tense are most determinate in the continuous/progressive sense, followed by the continuous/non-progressive, and least determinate in the habitual sense. In other words, "continuous/progressive" is more focal to imperfectivity than "continuous/non-progressive", which is in turn more focal than "habitual":

continuous/progressive < continuous/non-progressive < habitual

(b) judgments on sentences including the perfect tense are most determinate when the sentence refers to an action with no internal structure, followed by an action with specified duration, and least determinate when the sentence refers to an iterative situation where the frequency is specified.

5.3.4 Subjects

As in Pilot 2, subjects who participated in Pilot 3 were English-speaking learners of Italian at three levels of proficiency: (a) low intermediate, (b) advanced, and (c) near-native. A group of native Italian speakers were also tested.
Learners were students of Italian in the Italian Department at the University of Edinburgh. All had started learning Italian at the university, and none had Italian origins. As in Pilot 2, near-native speakers were selected through a preliminary interview: the criterion for inclusion in this group was native-like fluency, correctness, and lexical choice. Phonological factors in performance were not taken into account. The native speakers were Italians living in Edinburgh, either permanently or temporarily.
All subjects volunteered to take part in the experiment.
As in Pilot 2, the number of subjects in each group was as follows:

- 18 natives
- 6 near-natives

- 6 advanced learners (with at least three years of exposure to Italian, of which one spent in Italy)
- 6 low intermediate learners (with less than one year of exposure to Italian)

5.3.5 Materials, tasks and design

The test materials for Pilot 3 were analogous to the materials used in Pilot 2 (test materials can be found in Appendix A).
Subjects had to judge the acceptability of 90 sentences, divided into three sets. Each set consisted of (a) 15 grammatical sentences, of which 10 represented the categories of use for the imperfective area and 5 represented the categories of use for the perfective area, and (b) 15 sentences lexically identical to the ones above but differing in the use of tense, i.e. having the present perfect tense instead of the imperfect and vice versa. Although most of these sentences were ungrammatical, some of them were actually acceptable. For example, the first set contained both the sentence (a) and the corresponding sentence (b):

(a) Ieri ho dormito fino a mezzogiorno
'Yesterday I slept until noon'
(b) Ieri dormivo fino a mezzogiorno
'Yesterday I was sleeping until noon'

Each set therefore included 30 sentences. The three sets were equivalent to one another in that they consisted of different lexicalizations of the same categories. For example, the category exemplified by sentence (a) above in set 1 (Perfective: end-point of situation specified) was also represented by the following sentences in sets 2 and 3 respectively:

(c) Ieri abbiamo guardato la TV fino a mezzanotte
'Yesterday we watched TV until midnight'
(d) Ieri siamo rimasti al ristorante fino all'una
'Yesterday we stayed at the restaurant until one o'clock'

As before, the three sets will henceforth be referred to as **Lex 1**, **Lex 2** and **Lex 3**. Sentences within each set were presented in different random orders.

The research design was the same Latin square design as in Pilot 2. Subjects were asked to perform in the same three tasks: Magnitude Estimation, cards, and Ranking (see p. xx for a description).

In order to identify an acceptability hierarchy in the judgments of native speakers, the following steps were taken:

(a) A s in Pilot 1 and Pilot 2, the differences were calculated, for each subject and each sentence type, between responses on correct tenses and those on incorrect tenses. Arithmetic means were then computed on the resulting "preference scores" for Ranking and Card-Sorting; geometric means for Magnitude Estimation.

(b) a series of repeated-measures ANOVAs were performed for each task separately:
   i. a two-way ANOVA (proficiency level x semantic category) , on all data;
   ii. two separate two-way ANOVAs (proficiency level x semantic category) for imperfectivity and perfectivity.

(c) post-hoc Tukey tests of significance were performed on the means. The following comparisons were made:
   i. among categories within each proficiency level
   ii. among proficiency levels for each category.

The critical values of q for all comparisons for each method are as follows:

Imperfectivity

Magnitude Estimation	4.43
Card-Sorting	42.08
Ranking	63.56

Perfectivity

Magnitude Estimation	4.43
Card-Sorting	42.3
Ranking	117.59

5.3.6 Results

The presentation of the results will be divided into two sections. The first one will be concerned with the methodology of the research, whereas the second one will examine its theoretical implications.

The mean acceptability judgments are reported in Tables 5.6, 5.7, and 5.8 corresponding to Magnitude Estimation, Card-Sorting, and Ranking. As before, the size of the numbers indicates the strength of preference for a given tense, and the sign expresses preference for the correct (positive sign) or for the incorrect (negative sign) tense.

TABLE 5.8. Pilot 3: mean tense preferences (in logarithmic form) in **Magnitude Estimation** judgments on imperfective vs perfective (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

	INT	ADV	NNS	NS
Imperfectivity				
TEMPORAL PHRASE	0.024	0.015	0.078	0.134
TEMPORAL CLAUSE	0.400	0.393	0.149	0.536
STARTING POINT NAMED	0.359	0.651	0.462	0.625
STARTING POINT INFERRED	-0.080	0.281	0.455	0.566
SERIES OF EVENTS	0.013	-0.017	0.025	0.002
STATIC	0.335	0.155	0.126	0.124
MODAL	-0.017	0.317	0.092	0.597
HABIT IN THE PAST	0.329	0.157	0.241	0.398
ONGOING SITUATION	-0.075	0.06	0.038	0.159
ITERATIVE SITUATION	0.262	0.146	-0.048	-0.044
Perfectivity				
DURATION SPECIFIED	0.223	0.279	0.298	0.589
END-POINT SPECIFIED	0.234	0.614	0.194	0.539
NO INTERNAL STRUCTURE	0.496	0.526	0.124	0.489
ITERATIVE/FREQUENCY SPECIFIED	-0.004	0.196	0.257	0.518
ITERATIVE/DURATION SPECIFIED	-0.091	0.032	0.173	0.399

TABLE 5.9. Pilot 3: mean tense preferences in **Card-Sorting** judgments on imperfective vs perfective (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

	INT	ADV	NNS	NS
Imperfectivity				
TEMPORAL PHRASE	2.167	2.833	0.833	1.722
TEMPORAL CLAUSE	3.333	7.833	2.833	4.000
STARTING POINT NAMED	1.000	1.333	3.000	5.278
STARTING POINT INFERRED	0.500	1.000	3.333	4.000
SERIES OF EVENTS	0.167	-0.167	1.000	1.000
STATIC	1.000	4.000	2.667	2.278
MODAL	-0.667	-0.333	2.667	4.222
HABIT IN THE PAST	1.333	2.500	1.667	3.444
ONGOING SITUATION	1.167	4.167	1.167	1.278
ITERATIVE SITUATION	2.333	4.167	-0.167	-0.556
Perfectivity				
DURATION SPECIFIED	-0.667	2.167	2.333	4.722
END-POINT SPECIFIED	1.000	4.167	3.667	4.389
NO INTERNAL STRUCTURE	3.167	7.500	3.667	4.389
ITERATIVE/FREQUENCY SPECIFIED	1.500	-0.500	4.000	4.056
ITERATIVE/DURATION SPECIFIED	-1.167	0.833	-0.500	2.167

TABLE 5.10. Pilot 3: mean tense preferences (in logarithmic form) in **Ranking** judgments on imperfective vs perfective (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

	INT	ADV	NNS	NS
Imperfectivity				
TEMPORAL PHRASE	2.833	2.333	1.333	1.500
TEMPORAL CLAUSE	7.833	4.333	1.333	7.056
STARTING POINT NAMED	1.333	5.667	3.833	7.222
STARTING POINT INFERRED	1.000	5.167	3.667	4.722
SERIES OF EVENTS	-0.167	0.500	0.833	-0.111
STATIC	4.000	3.000	1.667	5.222
MODAL	-0.333	2.333	2.167	6.833
HABIT IN THE PAST	2.500	1.667	1.833	6.056
ONGOING SITUATION	4.167	1.667	0.000	3.500
ITERATIVE SITUATION	4.167	2.333	-0.667	1.500
Perfectivity				
DURATION SPECIFIED	2.167	1.167	1.000	7.444
END-POINT SPECIFIED	4.167	4.167	2.833	7.111
NO INTERNAL STRUCTURE	7.500	5.167	3.333	7.389
ITERATIVE/FREQUENCY SPECIFIED	-0.500	2.667	2.500	6.389
ITERATIVE/DURATION SPECIFIED	0.833	-1.500	0.833	4.167

A graphic display of these means can be seen in Fig. 5.6a, b (Magnitude Estimation), Fig. 5.7a, b (Card-Sorting) and Fig. 5.8 a, b (Ranking). The (a) graphs concern judgments on imperfective forms and the (b) graphs concern perfective forms.

FIGURE 5.6a. Pilot 3: mean tense preferences in **Magnitude Estimation** judgments on **imperfective** sentences (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

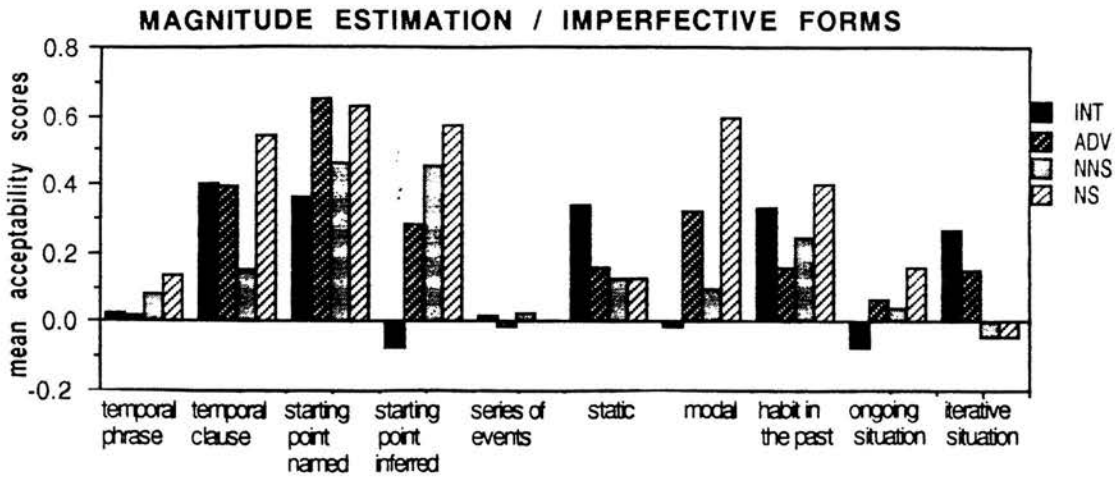


FIGURE 5.6b. Pilot 3: mean tense preferences in **Magnitude Estimation** judgments on **perfective** sentences (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

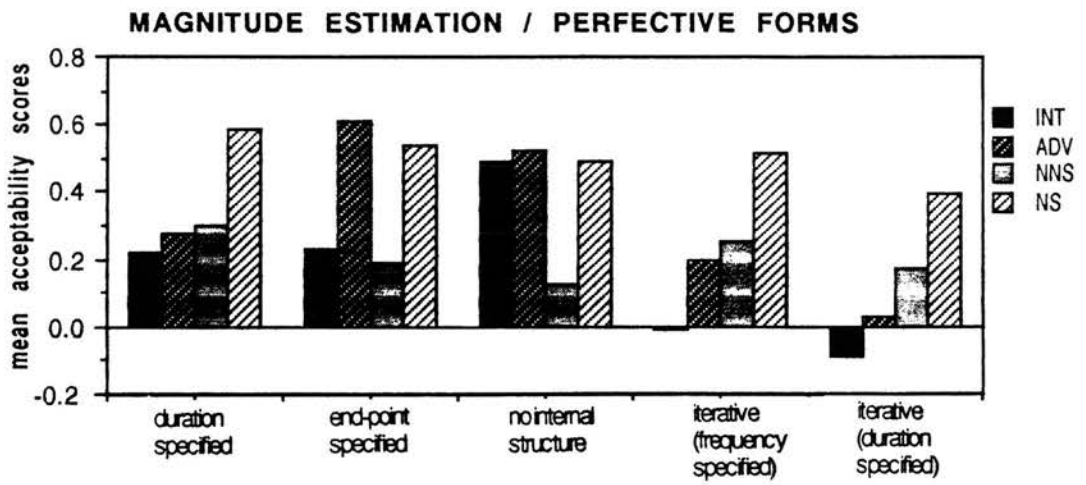


FIGURE 5.7a. Pilot 3: mean tense preferences in **Card-Sorting** judgments on **imperfective** sentences (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

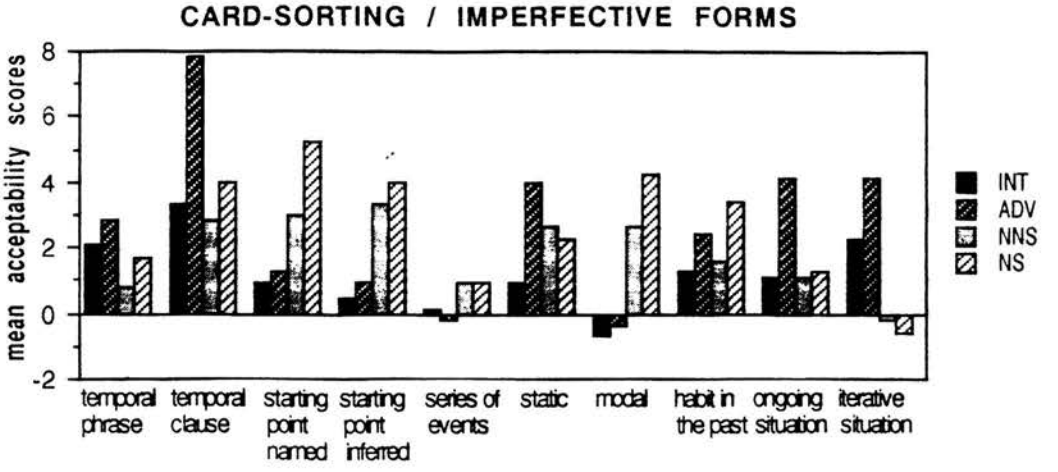


FIGURE 5.7b. Pilot 3: mean tense preferences in **Card-Sorting** judgments on **perfective** sentences (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

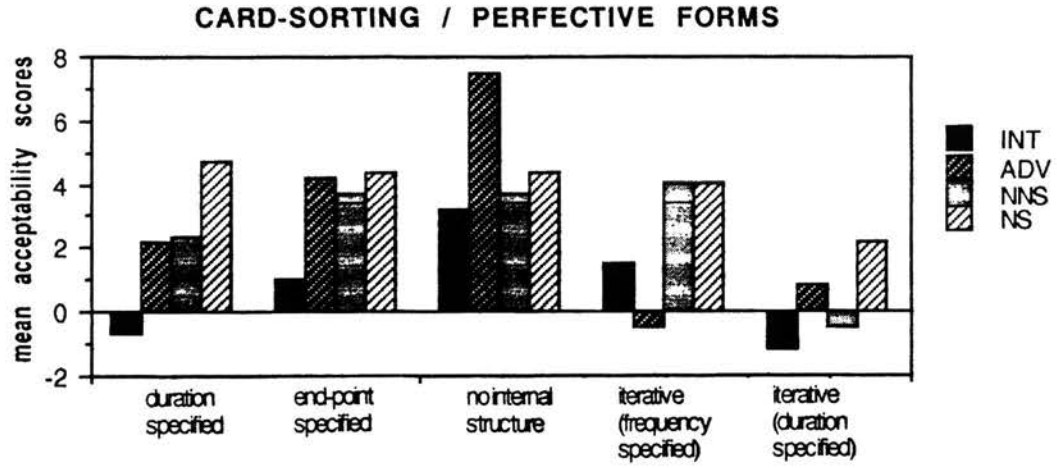


FIGURE 5.8a. Pilot 3: mean tense preferences in **Ranking** judgments on **imperfective** sentences (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)

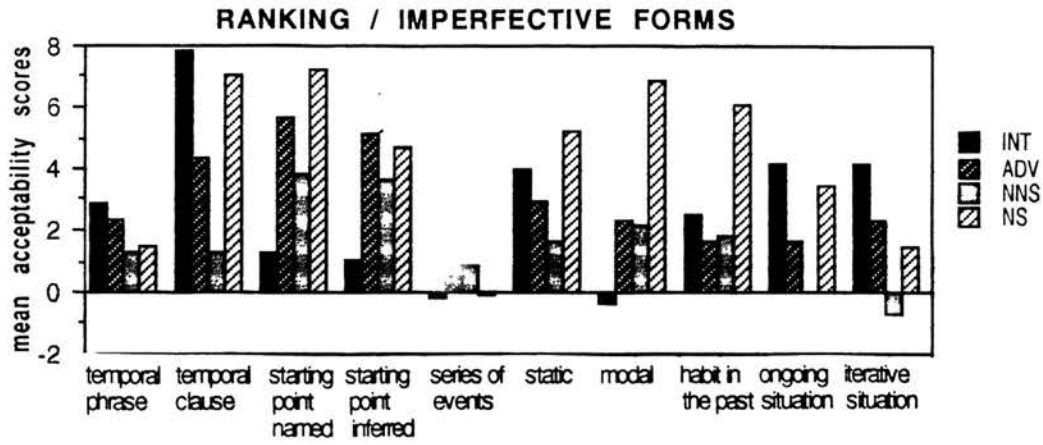
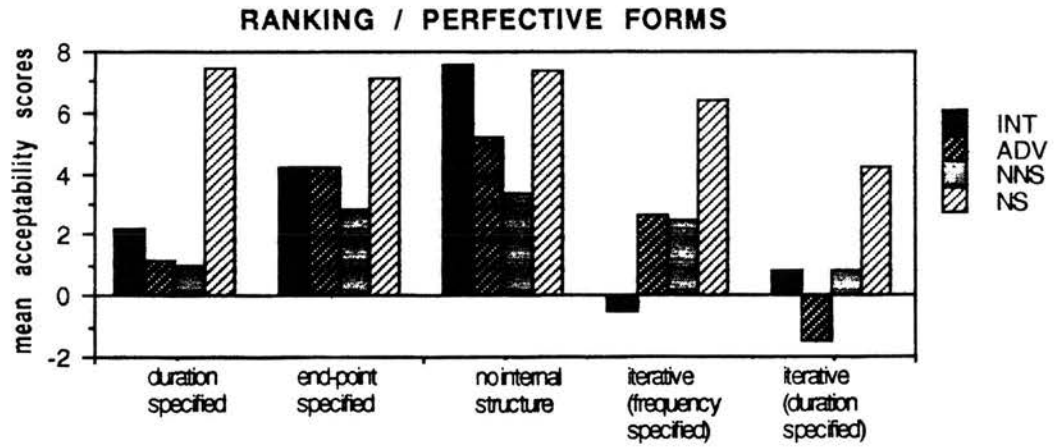


FIGURE 5.8b. Pilot 3: mean tense preferences in **Ranking** judgments on **perfective** sentences (INT = intermediate; ADV = advanced; NNS = near-native speakers; NS = native speakers)



5.3.6.1 Imperfective forms

The judgments on imperfective forms obtained through the three methods are clearly sensitive to differences among aspectual categories. Statistically, this is reflected by a significant main effect of semantic category in all the three ANOVAs: Magnitude Estimation, $F(9, 288) = 9.48$, $p < .0001$, Card-Sorting, $F(9, 288) = 4.12$, $p < .0001$, Ranking, $F(9, 288) = 3.3$, $p < .0008$. Judgments are also differentiated according to proficiency level: both Magnitude Estimation and Card-Sorting (but not Ranking) show a significant interaction of proficiency level and semantic category ($F(27, 288) = 2.27$, $p < .005$ and $F(27, 288) = 2.88$, $p < .0001$, respectively). Only Magnitude Estimation gives a significant main effect of proficiency level, $F(3, 32) = 3.44$, $p < .03$.

Inspection of the means for native speakers' judgments in Table 5.6 - 5.8 indicates that there is broad consistency among the three methods in producing the following rank ordered groups of aspectual categories:

TABLE 5.11. Pilot 3. Imperfective sentences: rank-ordered aspectual categories (overall results)

starting point named
non-progressive (modal)
starting point inferred
point of reference (clause) along time line
situation habitual of whole past period
situation characteristic of limited past period
point of reference (phrase) along time line
non-progressive (static)
series of events described as single whole
iterative situation (frequency/time limits unspecified)

The twelve categories have been divided into three blocks: this means that, while there is variation among the orders of individual categories within blocks, the rank order of the blocks is largely invariant across methods. However, Tukey tests of significance applied to the means reveal that the three methods differ in

terms of sensitivity to degrees of acceptability: Magnitude Estimation is more capable of capturing differences among aspectual categories.
Let us consider native judgments first. In the Magnitude Estimation data from native speakers, preferences for the imperfect tense in 'starting point named' sentences are significantly stronger than those for 'iterative' sentences (q = 4.99), and for 'series of events' sentences (q = 4.64). The use of the imperfect in 'iterative' sentences is also significantly less acceptable than its use in 'starting point inferred' sentences (q = 4.55) and in 'non-progressive/modal' sentences (q = 4.78). Similarly, preferences for the imperfect in 'series of event' sentences is weaker than in 'non-progressive/modal' sentences (q = 4.43).

The Card-Sorting data from native speakers only give a significant difference between 'starting point named' and 'iterative situation' sentences. The Ranking data do not show any significant difference among tense preferences for aspectual categories.

The patterns of non-native judgments show that in most cases non-native acceptability values remain quite distant from native values, even though at the most advanced level they tend to fit the same hierarchies. The three methods reveal very few significant differences in terms of tense preferences. In the Magnitude Estimation data from near-native speakers, the imperfect with 'starting point named' sentences is preferred to its use in 'iterative situation' sentences (q = 4.44). In the advanced learners' judgments, there are differences between 'starting point named' sentences and both 'point of reference (phrase)' (q = 4.74) and 'series of events' sentences (q = 4.98). No significant differences are found at the intermediate level.

In the Card-Sorting data, the only significant differences are at the advanced level between 'point of reference (clause)', on the one hand, and 'series of events' sentences (q = 59.69), as well as 'non-progressive/static' (q = 60.94) on the other. No significant differences are found at either the intermediate or the near-native levels.

Finally, the Ranking method does not produce any significant differences among aspectual categories.

From a developmental perspective, the few significant differences among proficiency levels are given only by Magnitude Estimation, and concern intermediate learners and native speakers. Native tense preferences are significantly stronger than intermediate learners' with respect to 'starting point inferred' sentences (q = 4.81) and 'non-progressive/modal' sentences (q = 4.57).

5.3.6.2 Perfective forms

Perfectivity appears to be a more homogeneous area in which native speakers do not make sharp distinctions among aspectual categories. It was predicted that the aspectual situations that unambiguously require the use of the present perfect tense (i.e. duration specified, end-point specified, and no internal structure), would be more determinate than the remaining two categories, defining iterative situations in which either the frequency is specified or the temporal markers refer to the whole period. What results from the three methods is a loose rank order, in which only the extremes are constant, but not the middle categories:

TABLE 5.12. Pilot 3. Perfective sentences: rank-ordered aspectual categories (overall results)

duration specified
end-point of situation specified
situation without internal structure
iterative situation (frequency/no of times specified)
iterative situation (temporal markers referring to it as a whole)

The ANOVA gives a significant main effect of proficiency level only for Magnitude Estimation, $F(3, 32) = 4.86$, $p < .007$, and for Ranking, $F(3, 32) = 4.15$, $p < .01$. It also gives a significant effect of semantic category for all three methods: Magnitude Estimation, $F(4, 128) = 4.66$, $p < .002$; Cards, $F(4, 128) = 9.39$, $p < .0001$, and Ranking, $F(4, 128) = 5.59$, $p < .0004$. There is a significant interaction of semantic category and proficiency level but only in the Cards data, $F(12, 128) = 2.19$, $p < .02$.

Post-hoc Tukey tests, however, do not give any significant differences, either within- or across levels, for any of the three methods.

5.3.7 Problems

How does the obtained order for imperfectivity fit the predictions? It will be recalled that continuous/progressive categories were predicted to be more determinate than continuous/non-progressive, and habitual categories were predicted to be the least determinate. The prediction is only partly fulfilled. While three continuous/progressive categories (starting point named; starting point inferred; point of reference (clause) along a time line) are at in the top block of the hierarchy, a fourth (point of reference(phrase)) is at a much lower place. Furthermore, a non-progressive category (modal) is one of the two most determinate ones. Two of the habitual categories are consistently lowest in terms of determinacy (ongoing situation in limited period; iterative situation), while a third (situation characteristic of whole past period) is at a higher position.

A possible explanation is that there are problems underlying the use of acceptability judgments in this domain. Consider sentences involving a temporal phrase as a point of reference, such as

(33) a. Ieri alle due ascoltavo musica classica
'Yesterday at two o'clock I was listening to classical music'

b. Ieri alle due scrivevo delle lettere
'Yesterday at two I was writing some letters'

These sentences are ambiguous. The ambiguity lies in the fact that the temporal phrase may be perceived as the starting point of the action rather than a point of reference along a time line (Yesterday at two *I started to write* some letters) and thus become a case for the use of the perfect tense. This may explain the difference between temporal phrases and temporal clauses (the latter being more difficult to interpret as starting points), as well as the low position on the hierarchy of this category .

The category 'situation characteristic of whole past period' is also potentially ambiguous in that it may or may not denote completeness, involving two different readings. Thus, the (isolated) sentence (34) could take either tense:

(34) a. In quel periodo Mario portava la barba
b. in quel periodo Mario ha portato la barba
'At that time Mario had a beard'

The ambiguity, however, could be disambiguated by a following context such as "...e la porta ancora" (and he still has one), which would be possible with (34)a, but not with (34)b.

In general, therefore, it became apparent from the results of Pilot 3 that many of the sentences exemplifying aspectual categories within the imperfective domain could be affected by varying the context. Given that sentences were presented in isolation, it is plausible to assume that in some cases the informants themselves provided the missing context, thus facilitating the task of expressing acceptability judgments.

Furthermore, the inherent ambiguity of some of the test sentences may be responsible for the apparent lack of development in non-native linguistic intuitions: what distinguishes natives from non-natives in this study may be not so much their ability to discriminate among aspectual categories as their ability to provide 'on-line' contextualizations (see Chapter 1, section 1.1.1.1).

5.3.8 Conclusions

The results of Pilot 3 were informative in two respects:

(a) although it showed differences in the way imperfective constructions are perceived, these differences do not seem to be scalable along an acceptability hierarchy. This is plausibly due to a problem inherent in the use of decontextualized acceptability judgments in this area of grammar. For certain categories of sentences presented in isolation, there could be no principled way of establishing whether variation in acceptability is due to grammatical or to contextual factors.

(b) methodologically, the study once again confirmed the viability of Magnitude Estimation as a method for the elicitation of acceptability judgments, and its comparatively higher sensitivity to differences in degrees of acceptability.

To conclude, the pilot studies described in this chapter encouraged us to pursue both the methodological and the theoretical lines of investigations that motivated our research. They

were useful in allowing a reformulation of the research questions and a series of modifications of the experimental materials and design that were employed in the main study. These will be outlined in Chapter 6.

CHAPTER 6
THE MAIN STUDY

6.0 Introduction

This chapter is concerned with the structure of the main study. First, it will outline the decisions that were taken on the basis of the results of the pilot studies described in Chapter 5. It will then provide details of the experimental materials, design, subjects, and method employed in the main study.

6.1 Post-pilot developments

On the basis of the results obtained in the pilot studies, three main decisions were taken.

First, it was decided to focus exclusively on the area of auxiliary selection. This domain of grammar proves to be less sensitive to contextual variation, arguably because it is governed by an interplay of syntactic and semantic factors (as opposed to aspect, which is characterized mainly by semantic features). This decision was supported by the results of the pilot studies: Pilot 1 and 2 were encouraging in showing systematic variation among intransitive verbs, and Pilot 2 provided preliminary evidence that non-native acceptability values develop in the direction of native values.

Second, the set of sentence types for auxiliary assignment was expanded to include a number of syntactic properties related to unaccusativity: choice of auxiliary, *ne*-cliticization with both unergative and unaccusative verbs; restructuring constructions with related phenomena such as clitic-climbing, auxiliary change, impossibility of clefting. Transitive verbs were excluded, and the main focus was exclusively centred on the properties of intransitive (i.e. unaccusative and unergative) verbs.

Third, the Ranking measurement was discarded, whereas Card-Sorting and Magnitude Estimation were retained. The findings of the pilot studies indicate a difference between Magnitude Estimation and Card-Sorting on the one hand, and Ranking on the other. In the pilot studies, Magnitude Estimation provided a more fine-grained picture of acceptability in native judgments and of development in non-

native judgments. Furthermore, the Ranking test has lower face validity than Card-Sorting. Given that one of the aims of this study is to test whether Magnitude Estimation works as a technique for the elicitation of acceptability judgments, it was decided that Card-Sorting would be a good point of comparison as representative of the set of more ordinary methods that make use of successive categories (as Chapter 2 showed in some detail).

Fourth, it was decided to test a representative sample of French near-native speakers of Italian in addition to the English-speaking learners. Pilot 1 suggested that the differences between the auxiliary systems of Italian and French are responsible for the directionality of difficulty in second language acquisition: French learners appear to start from a favourable base in the task of learning the Italian system. But how does French compare to English with respect to auxiliary selection? One can say that:

(a) French has a system of auxiliary selection that is in parametric variation with the Italian system. Its class of unaccusative verbs coincides with the Italian class and exhibits similar syntactic behaviour. However, French requires ETRE almost exclusively for verbs belonging to the two core categories at the top of the Unaccusative Hierarchy, and AVOIR for all the other verbs that select ESSERE in Italian. French has clitics, but no restructuring constructions, although it used to have them at a previous stage of its historical evolution. One can therefore argue that French presents *partial instantiation* of the properties exhibited by Italian.

(b) English has a semantic class of unaccusative verbs but does not instantiate any of the syntactic properties under investigation. It has no system of auxiliary selection and no clitics. One can therefore say that English shows virtually *no instantiation* of the properties in question.

The interesting question therefore arises as to whether 'partial instantiation' is a better starting point than 'no instantiation' for the purpose of acquiring the Italian auxiliary system. In order to address this question, it was decided to test French native speakers at the highest level of proficiency in Italian as a foreign language.

6.2 Materials

The test materials employed in the main study included sentences on three distinct areas of the Italian grammar:

(a) unergative verbs
(b) unaccusative verbs
(c) restructuring constructions

Each subset comprised sentences that exemplified the semantic or syntactic categories subdividing that verb class. These categories (with relevant verbs) are listed below:

(a) unergative
- motional (*nuotare, viaggiare, camminare, passeggiare*)
- non-motional (*giocare, dormire, lavorare, parlare*)
- +unaccusative alternant (*correre, saltare, volare, rotolare*)

(b) unaccusative
- change of location (*venire, arrivare, ritornare, cadere*)
- continuation of state (*restare, sopravvivere, rimanere, durare*)
- existence of state (*esistere, appartenere, sembrare, piacere*)
- +transitive alternant (*aumentare, affondare, cambiare, migliorare*)
- +unergative alternant (*correre, saltare, volare, rotolare*)

(c) restructuring
- Raising, non-restructured (*potere, dovere, continuare*)
- Raising, restructured
- Control, restructured (*volere, sapere, cominciare*)
- Control, restructured

In the unergative and the unaccusative sets, each semantic/syntactic category was represented by:

(i) sentences in basic word order (i.e. declarative sentences with full NP arguments); examples:
Paola ha passeggiato nel giardino con i bambini (unergative, motional)
Francesca è arrivata in ritardo (unaccusative, change-of-state)

(ii) sentences presenting NE-cliticization; examples:
**Di ragazzi, ne hanno lavorato molti in questo ristorante* (unergative, non-motional)
Di stranieri in Italia, ne sono venuti molti l'anno scorso (unaccusative, change-of-state)

Each of these word order types was represented by a two sentences: one with ESSERE and one with AVERE.

In the restructuring set, each syntactic category was represented by four sentence types:

(i) sentences in basic word order (i.e. declarative sentences with full NP arguments); examples:
Giovanna è dovuta rimanere a Roma (Raising, restructured)
Maria non ha potuto andare alla festa (Control, non-restructured)

(ii) sentences presenting clefting; examples:
E' tornare in albergo che la famosa attrice ha voluto (Control, non-restructured)
**E' rimanere a Roma che Giovanna è dovuta* (Raising, restructured)

(iii) sentences presenting clitic-climbing; examples:
Al concerto del lunedì, la gente ci è continuata ad andare (Raising, restructured)
**Alla festa, Maria non ci ha potuto andare* (Control, non-restructured)

(iv) sentences presenting post-verbal pronominal arguments, with no clitic movement; examples:
A casa, Francesca è dovuta rimanerci (Raising, restructured)
A scuola, mia figlia non ha potuto andarci (Control, non-restructured)

6.2.1 Test categories

The categories used in the judgment test are listed in Table 6.1.

TABLE 6.1: Verb categories used in the main study

1. unergative, motional, basic, ESSERE
2. unergative, motional, basic, AVERE
3. unergative, motional, +Ne-cliticization, ESSERE
4. unergative, motional, +Ne-cliticization, AVERE
5. unergative, non-motional, basic, ESSERE
6. unergative, non-motional, basic, AVERE
7. unergative, non-motional, + Ne- cliticization, ESSERE
8. unergative, non-motional, + Ne- cliticization, AVERE
9. unergative, with unaccusative alternant, basic, ESSERE
10. unergative, with unaccusative alternant, basic, AVERE
11. unergative, with unaccusative alternant, +Ne-cliticization,ESSERE
12. unergative, with unaccusative alternant, +Ne-cliticization, AVERE

13. unaccusative, change of location, basic, ESSERE
14. unaccusative, change of location,, basic, AVERE
15. unaccusative, change of location, +Ne-cliticization, ESSERE
16. unaccusative, change of location, +Ne-cliticization, AVERE
17. unaccusative, continuation of state, basic, ESSERE
18. unaccusative, continuation of state, basic, AVERE
19. unaccusative, continuation of state, +Ne-cliticization, ESSERE
20. unaccusative, continuation of state, +Ne-cliticization, AVERE
21. unaccusative, existence of state, basic, ESSERE
22. unaccusative, existence of state, basic, AVERE
23. unaccusative, existence of state, +Ne-cliticization, ESSERE
24. unaccusative, existence of state, +Ne-cliticization, AVERE
25. unaccusative, with transitive alternant, basic, ESSERE
26. unaccusative, with transitive alternant, basic, AVERE
27. unaccusative, with transitive alternant, +Ne-cliticization, ESSERE
28. unaccusative, with transitive alternant, +Ne-cliticization, AVERE
29. unaccusative, with unergative alternant, basic, ESSERE
30. unaccusative, with unergative alternant, basic, AVERE
31. unaccusative, with unergative alternant, +Ne-cliticization,ESSERE
32. unaccusative, with unergative alternant, +Ne-cliticization, AVERE

33. Raising, non-restructured, basic, AVERE
34. Raising, non-restructured, clefting, AVERE
35. Raising, non-restructured, clitic-climbing, AVERE
36. Raising, non-restructured, no clitic movement, AVERE
37. Raising, restructured, basic, ESSERE
38. Raising, restructured, clefting, ESSERE
39. Raising, restructured, clitic-climbing, ESSERE
40. Raising, restructured, no clitic movement, ESSERE
41. Control, non-restructured, basic, AVERE
42. Control, non-restructured, clefting, AVERE
43. Control, non-restructured, clitic-climbing, AVERE
44. Control, non-restructured, no clitic movement, AVERE
45. Control, restructured, basic, ESSERE
46. Control, restructured, clefting, ESSERE
47. Control, restructured, clitic-climbing, ESSERE
48. Control, restructured, no clitic movement, ESSERE

6.2.2 Preparation of test materials

In the construction of test sentences, attention was paid to factors such as lexical complexity and length. The lexical items chosen for inclusion in the sentences were all high-frequency words that students of Italian are likely to encounter at the beginning of a course. Basic sentences were similar in length, although sentences exemplifying different word orders were slightly longer. This is due to the fact that constructions including clitics (such as NE-cliticization, clitic-climbing, and no-clitic-movement sentences) required a minimum amount of context in order to make sense. Given the practical difficulty involved in presenting fully contextualized sentences, an element of context was built into the test sentences themselves, in the form of topicalized argument. So, for example, a sentence presenting an existence-of-state unaccusative verb with NE-cliticization was constructed with the inverted subject topicalized and fronted:

DI DINOSAURI, ne sono esistiti molti

Similarly, a sentence presenting clitic-climbing had a fronted topic:

ALLA TUA FESTA, Maria non ci è potuta andare

The inevitable shortcoming of this operation was the relative markedness and length of sentences containing clitics compared to basic sentences.

6.3 Tasks

Acceptability judgments were elicited by two methods:

(a) Magnitude Estimation
(b) Card-Sorting.

The Magnitude Estimation test involved the presentation of isolated sentences printed on slides. The sentences were shown to subjects on an overhead projector. A mask was moved down the slide, revealing a sentence at a time at intervals of about 5 seconds. Subjects were instructed to write their responses on individual answer sheets that had been provided beforehand. The answer sheets had a column of boxes printed in correspondence of sentences numbers (1 to 48). Subjects had to write each numerical response in the box corresponding to the sentence they had just been shown.

Instructions read as follows:

> "Isolated sentences will be projected on the screen in front of you, one at a time. You will have to judge the acceptability of these sentences. Give the first sentence any number you wish (but not 0); then assign successive sentences numbers that are proportional to the first number you chose. For instance, if you gave the number 6 to the first sentence and if you think that the second sentence is twice as acceptable as the first, give it the number 12. Similarly, if you think that the second sentence is only a third as acceptable as the first one, choose the number 2. If the second sentence is only a little less acceptable than the first, give it the number 5, and so on. You can assign the same number to different sentences if you think that they have the same degree of acceptability.'

Examples followed on the instruction sheet. This can be seen in Appendix A.

The Card-Sorting test consisted of the untimed arrangement of test sentences into piles. Subjects were given a set of 48 cards, each containing one sentence printed on one side, numbered from 1 to 48. They were invited to sit at a desk and to work on the distribution of sentences into piles, starting from the least acceptable ones on their left hand side and ending with the most acceptable ones on the right end side. Piles could contain any number of sentences, provided they were judged to have the same degree of acceptability. Once they had finished arranging the sentences, they were asked to record their responses on a sheet of graph paper, by writing the sentence numbers in each pile in separate columns, ordered like the piles on their desk.
The instructions for this task were as follows:

> "You have a set of cards, each with a sentence printed on it. Your task is to arrange the sentences in different piles, according to their degree of acceptability. You may form as many piles as you wish, even if a pile contains only one sentence. Once you have divided all the sentences into piles, arrange the piles across your desk, starting from 'least acceptable' on your left towards 'most acceptable' on your right. Then use the sheet of graph paper to record which sentences are in which pile and how the piles are arranged across the desk. Write columns of numbers corresponding to the sentences in each pile, again from left (= least acceptable) to right (= most acceptable)."

The procedure was then illustrated by a series of examples, that can be seen in Appendix A.

6.4 Subjects

Six groups of informants participated in the experiment, for a total of 180 subjects :

- 32 beginners (first-year university students with six months of exposure to Italian);
- 36 intermediate (second and third-year students with at least eighteen months of exposure to Italian);
- 32 advanced (fourth-year students with at least three years of exposure to Italian, one of which spent in Italy);
- 24 English-speaking near-natives (with at least five years of exposure to Italian);
- 20 French-speaking near-natives (with at least five years of exposure to Italian);
- 36 native Italian speakers .

Learners (beginners to advanced) were students of Italian in the Italian Departments at the University of Edinburgh, Glasgow , and Strathclyde. They were in the age range 18-25. All had English as their native language, and none had Italian origins. All subjects had started learning Italian at the university; students who had had previously studied Italian at school were not included in the testing groups.

The near-natives had started learning Italian after the age of 15 (range of age of beginning exposure to Italian: 18 to 27); none had Italian origins. They ranged between 23 and 46 years in age and had had an average of 9 years of exposure to Italian (range: 5-15 years). While all subjects had received some formal instruction in Italian in the past, most of them reported having acquired Italian by living in Italy for long periods of time and by being in contact with Italians. Near-natives were recruited through (a) personal acquaintances, and (b) posted advertisements. Given the lack of a standardized 'test of near-nativeness', they were asked to participate in an individual screening interview prior to the experiment. The criterion for near-nativeness was native-like performance from the point of view of fluency,accuracy and lexical choice (although phonological accuracy was not taken into account). Subjects were selected only if they were able to pass as native speakers for the first 10 minutes of the interview. English near-natives were recruited both in Scotland and in Italy.
Three French near-natives were tested in France, and the remaining 17 in Italy.
Most of the French near-natives were teachers of French as a foreign language and had been residing in Rome for at least five years.
All the native Italian speakers were resident in Edinburgh at the time of testing, either temporarily or on a permanent basis. They were aged 22 to 51, and came from a wide range of professional backgrounds.

Participation in the experiment was entirely voluntary.

None of the informants who took part in the main study had participated in the pilot studies described in Chapter 5.

6.5 Design and presentation

There were 48 sentence structures in all (see Fig. 6.1). Of these 48 basic sentence forms, half (constructed with the correct auxiliary) were fully acceptable and the other half (constructed with ESSERE instead of AVERE and vice versa) presented varying degrees of unacceptability. Each form was represented by 4 different lexicalizations (these can be found in Appendix A)

The resulting 192 sentences were divided into 4 groups, each of which included all the 48 basic structures. Sentences were presented in different random orders within each lexicalization. Irrelevant distractors were not included, on the assumption that the different syntactic constructions represented in the test would act as distractors with respect to one another.

Within each of the six experimental groups, subjects were divided into two subgroups of equal size, A and B. Both A and B were then split into halves, so that each half (A1, A2, B1, B2) was given the two tests in a different order. The two halves of both A and B also received the tests in different lexicalizations. Assignment of subgroups to tasks and lexicalization was by Latin square, as shown in Table 6.2:

TABLE 6.2. Experimental design of the main study

Group A	A1	A2
Cards= lex 1, 4	2 ME	1 Cards
ME=lex 2, 3	4 Cards	3 ME
Group B	B1	B2
Cards= lex 2, 3	2 Cards	1 ME
ME=lex 1, 4	4 ME	3 Cards

Over two tests, each subject therefore judged 96 sentences.

6.6 Procedure

The administration of the tests to the learners, and to most of the English near-natives and native Italians, took place on university premises. Informants were usually tested in pairs, where both informants had been assigned the same lexicalization and the same task order. They sat in a quiet room, equipped with an overhead projector and two desks facing away from one another.
The French near-natives were administered the tests at the schools were they worked, in a similar experimental setting. The remaining few subjects were tested individually, normally at their place of work.

Instructions were given in English, French or Italian according to the native language of informants. All informants were given an individual folder containing:

(a) a general description of the purpose of the experiment (on the front cover)
(b) detailed instructions for the Magnitude Estimation task and an answer sheet to record responses;
(c) detailed instructions for the Card-Sorting task and a sheet of graph paper to record their responses.
(d) an envelope containing a set of cards, numbered from 1 to 48.

(These materials can be found in Appendix B).

Subjects were given the folder with the instructions and were asked to read the cover first. Then, they were directed to the first set of instructions inside the folder. Subjects could read at their own pace, and ask clarification questions at any time. When they felt satisfied that they had understood what was required of them, subjects started the first task that they had been assigned. Once they had completed the first test, they could proceed to the second test. After they had completed the second test, they were asked to express informal comments on any aspect of the experiment.
Testing was anonymous. All test materials in each folder had been marked with an unique sequence of letters and numbers, identifying the subject and his/her group and cell in the research design.

6.7 Statistical analysis of the results

The following steps were taken in the analysis of data:

(1) individual subjects files were prepared for both Magnitude Estimation and Card-Sorting;

(2) ANOVA input matrices were created in order to run by subject and by materials analyses for both Magnitude Estimation and Card-Sorting data;

(3) the BMDP statistical package was used to run the following parallel ANOVA analyses for Magnitude Estimation and Card-Sorting and for each of the areas of grammar investigated (unergative verbs, unaccusative verbs, restructuring verbs):

- by subject, all levels;
- by materials, all levels;
- by subject, natives vs near-natives;
- by materials, natives vs near-natives.

The 'all levels' analyses were run only on the four English-speaking groups of learners of Italian (beginners, intermediate, advanced, near-natives) and the native Italian group: the French near-natives were excluded. The 'natives vs near-natives' analyses were run only on the French near-native speakers, the English near-native speakers, and the native Italian speakers.

(4) further ANOVAs were conducted to analyse the mean differences in judged acceptability between the correct and the incorrect auxiliary for each of the three areas of grammar investigated; these analyses were once again run both by subject and by materials, on both the data from all levels and those from native vs near-natives;

(5) Min F' values were calculated for each pair of significant F_1 and F_2 values;[1]

(6) post-hoc Tukey tests of significance on the differences between pairs of F_1 means were applied for significant effects of interest.

6.8 Presentation of the results

The report on the results obtained in the main study will be divided into three sections:

(a) the first section (Chapter 7) will focus on methodology: it will compare the data elicited by the two techniques of Magnitude Estimation and Card-Sorting;

(b) the second section (Chapter 8) will concentrate on determinacy/indeterminacy in acceptability judgments. It will examine the systematic variation in the judgments of native speakers about the range of phenomena related to auxiliary selection, and the development of knowledge about auxiliary selection in Italian from beginner to near-native levels;

(c) the third section (Chapter 9) will explore the significance of the data for the question of ultimate attainment: it will focus on the representation of knowledge about auxiliary selection at the near-native level in speakers from two different language backgrounds (English and French) and on the comparison of near-native and native knowledge representations.

Finally, general conclusions from the results will be drawn in Chapter 10.

CHAPTER 7
RESULTS (1): METHODOLOGICAL ASPECTS

7.0 Introduction

The focus of this chapter is on a comparison of the two experimental procedures (Magnitude Estimation and Card-Sorting) employed for the collection of acceptability judgments in the main study.

One of the purposes of this study was provide an extended test of the validity of Magnitude Estimation as a method for the investigation of variation in linguistic acceptability. The experimental questions raised at the end of Chapter 2 were the following:

(a) is Magnitude Estimation applicable to the elicitation of acceptability judgments?
(b) is it sensitive to variation in linguistic acceptability? Does what it detects make any sense?
(c) how does it compare - in dealing with variation in acceptability - to ranking measurement scales making use of successive categories, such as Card-Sorting?

In order to provide answers to the above questions, the results will be discussed in this chapter only from the point of view of methodological adequacy; their linguistic significance (with respect to the theoretical questions raised by Chapter 3) will be the topic of Chapters 8 and 9.

The chapter is divided into four sections:

(1) the first section comments on the administration of the two tasks and their face-validity;
(2) the second section presents separate descriptive statistics for Magnitude Estimation and Card-Sorting;
(3) the third section compares the two methods with respect to the results obtained in the ANOVAs and includes a summary of all the significant main effects and interactions;
(4) the fourth section discusses the adequacy of Magnitude Estimation in the light of the results obtained.

7.1 Administration of the tests and face-validity

All informants performed in both the Magnitude Estimation and the Card-Sorting tasks, in different orders according to the research design described in Chapter 6. Informal

comments were collected at the end of the experimental session. These comments provide a basis for evaluating the face-validity of the two methods.

In general, Magnitude Estimation was regarded as a 'difficult' test by many informants. Most informants were completely new to the task of making ratio judgments, particularly under time pressure, and reluctant to trust themselves as reliable judges. They also felt that Magnitude Estimation required them to perform arithmetic operations that, even if basic, interfered with their concentration on the acceptability of sentences. They felt more confident in the Card-Sorting task, which required them to simply rank order sentences, and allowed them unlimited time as well as the opportunity to revise their initial judgments.
Subjects who received Magnitude Estimation before Card-Sorting generally found the former more demanding than subjects who received the two tasks in the reverse order. This is because when Card-Sorting was administered first, it gave them practice in the concept of relative acceptability and some familiarity with the sentence types.

Despite the subjective difficulty of the task, all subjects gave numerical responses without any apparent trouble. Responses were written at the pace required; no one lost track of the sequence of sentences shown on the screen, or of the corresponding boxes on their answer sheet.

Though psychologically 'easier' for many informants, Card-Sorting presented its own difficulties. The main one was indecisiveness: some took a long time to decide on the inclusion of some sentences in a particular pile, or to transfer sentences from one pile to the other. While Magnitude Estimation elicited first reactions to sentences, Card-Sorting elicited second (or third) reactions in a great number of cases. In general, the degree of indecisiveness and the number of modified choices grew proportionally with the level of proficiency of the informants. Some native and near-native speakers spent a long time over the final arrangement of their sentences into piles.
The administration of the Card-Sorting task did not present any other particular problem. Informants found the second part of the task (involving the transcription of the sentence numbers on paper) more cumbersome than the first one. Some informants spontaneously arranged the columns of numbers on an interval scale on the graph paper, reflecting their perception of the distance in acceptability between piles. The majority, however, simply put equal spacing between columns. Since this spacing information was not provided consistently, it was not taken into account in the analysis of the results.

7.2 Descriptive statistics

7.2.1 All subjects

The basic descriptive statistics for the two methods are reported in Table 7.1 and 7.2 below:

TABLE 7.1. Magnitude Estimation: basic descriptive statistics (raw data)

	MEAN	S.D.	RANGE	MODE
BEGINNERS	6.03	4.1	29	10
INTERMEDIATE	5.71	2.99	19	10
ADVANCED	6.28	3.59	24	10
ENGLISH NEAR-NATIVE	5.82	2.99	22	10
FRENCH NEAR-NATIVE	5.24	4.28	19	10
NATIVE SPEAKERS	5.09	5.53	49	10

TABLE 7.2. Card-Sorting: basic descriptive statistics

	MEAN	S.D.	RANGE	MODE
BEGINNERS	3.47	1.88	8	4
INTERMEDIATE	3.12	1.56	7	4
ADVANCED	3.2	1.68	8	4
ENGLISH NEAR-NATIVE	3.78	2.23	8	1
FRENCH NEAR-NATIVE	3.31	1.87	7	1
NATIVE SPEAKERS	3.72	2.65	18	1

The tables indicate that there are some important differences between the two methods. First of all, the ranges of scores used by subjects in the two different conditions appears to be much wider for Magnitude Estimation than for Card-Sorting. This was expected: in the Magnitude Estimation condition, informants had the whole number system at their disposal, whereas in Card-Sorting the number of piles they could make was constrained by the length of the desk and by the narrow resolution of this method. However, the means and the modes for both Magnitude Estimation and Card-Sorting suggest that the distribution of scores tends to concentrate on a narrower portion of the range. This will become obvious below in an analysis of scores by level.

The second observation suggested by Tables 7.1 and 7.2 is that there seems to be no correlation between the width of the score range used and the proficiency level of informants. The only distinction is between native and non-native informants: native speakers (as a group) use a wider range of numbers, although their group mean and mode are not higher than those of non-native speakers.

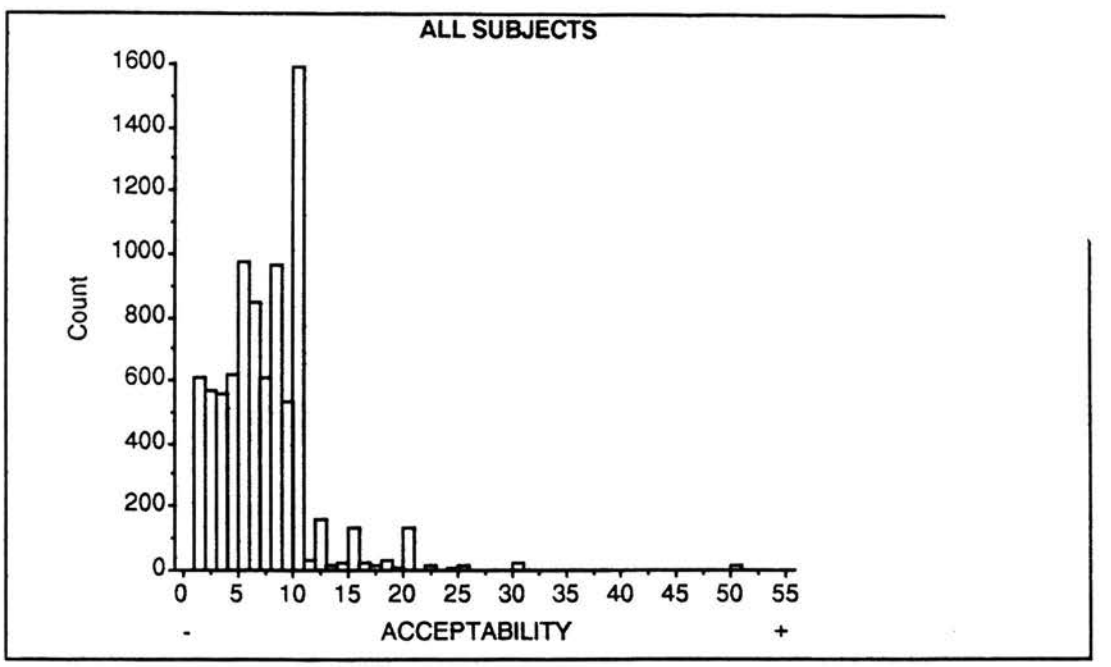
Third, while the mode for Magnitude Estimation scores is 10 at all levels, the mode for Card-Sorting is 4 for informants up to the advanced level, but it becomes 1 at the near-native and native levels. This suggests that, at least for Card-Sorting, learners have a tendency to give high acceptability ratings to sentences, but proficient speakers have the reverse tendency: they tend to assign low acceptability ratings.

7.2.1.1 Magnitude Estimation: score distribution

Let us now turn to a more detailed analysis of the overall frequency distribution of scores for the two methods. We will then look at the score distribution broken down by proficiency level.

The overall frequency distribution of the Magnitude Estimation scores is shown in Fig. 7.1.

FIGURE 7.1. Magnitude Estimation: frequency distribution of scores for all levels



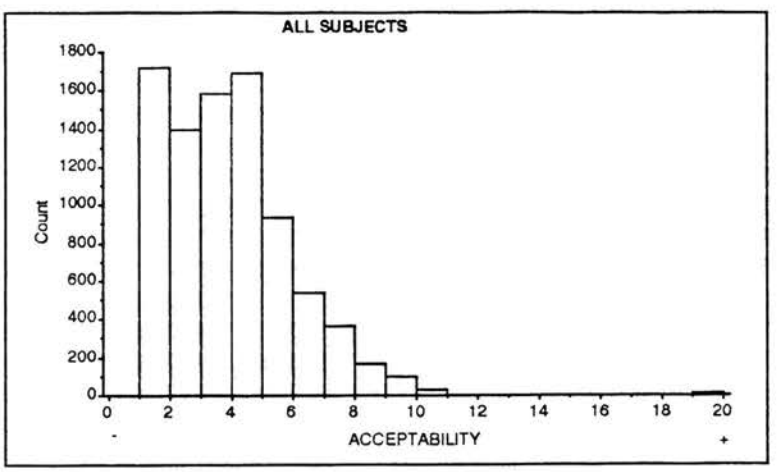
As shown by Fig. 7.1, the distribution of the Magnitude Estimation scores is heavily skewed: 92 % of all scores are between 1 and 10, whereas the remaining 8 % are between 11 and 50. The mode is the number 10, representing 19 % of the scores. The numbers between 1 and 9 are used in approximately equal proportion. This clearly suggests that most informants used a scale from 1 to 10.

7.2.1.2 Card-Sorting: score distribution

Fig. 7.2 reports theoverall distribution of scores for Card-Sorting.

FIGURE 7.2. Card-Sorting: frequency distribution of scores for all levels

---



The score distribution of Card-Sorting if also skewed: 75% of the scores lie between 1 and 4, and only 25 % between 5 and 20. This indicates that the majority of subjects sorted test sentences into four piles or, in other words, saw only four degrees of acceptability.

7.2.1.3 Summary of findings

The overall frequencies indicate that informants used a wider range of scores with Magnitude Estimation than with Card-Sorting, although not as wide a range as it was theoretically possible for them to use. The bulk of Magnitude Estimation scores gives a 10 to 1 range of acceptability; overall, this method gives a 50 to 1 range. The bulk of Card-Sorting, on the other hand, produces a 4 to 1 range, and overall a 20 to 1 range. Thus, Magnitude Estimation has been used more sensitively by the informants than Card-Sorting. We will come back to this issue in the concluding section of this chapter.

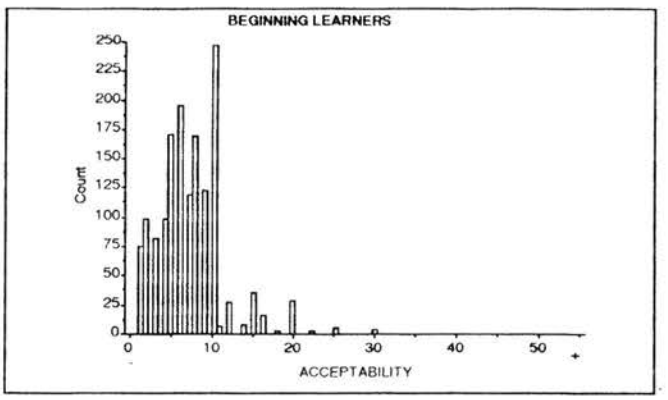
7.2.2 Descriptive statistics: by level

7.2.2.1 Magnitude Estimation

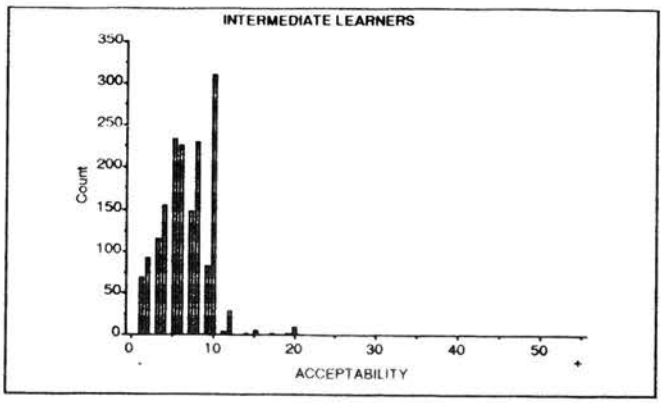
Figs. 7.3a-f below display the graphic representation of score distributions for each level.

FIGURE 7.3. Magnitude Estimation: frequency distribution of scores for Beginners (a), Intermediate (b), Advanced (c), English near-native (d), French near-native (e), and Native speakers (f)

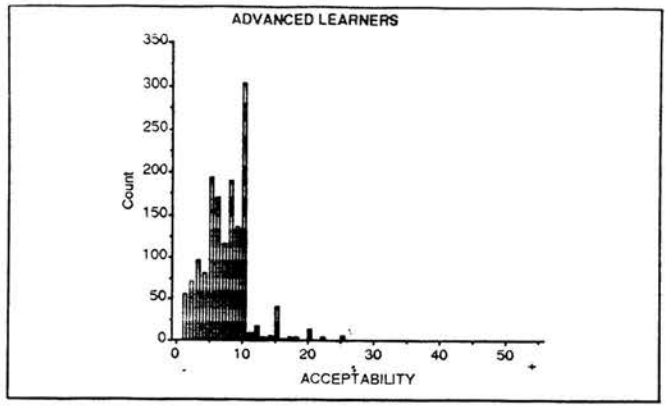
(a)



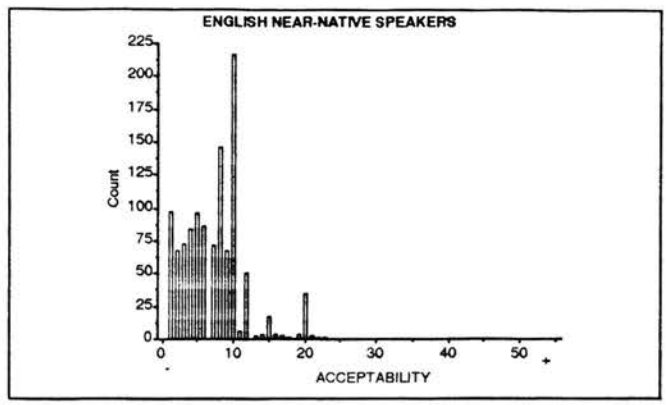
(b)



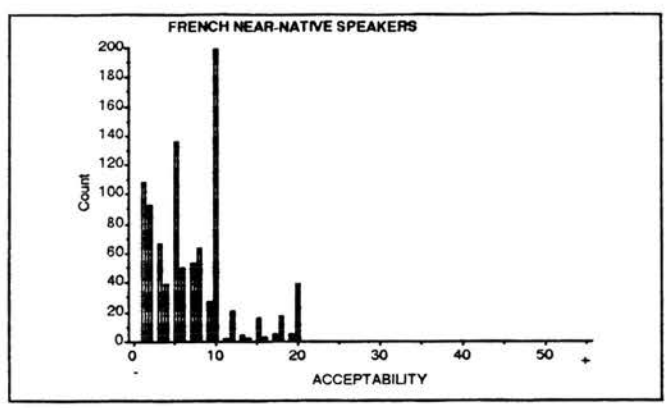
(c)


ADVANCED LEARNERS
Count
ACCEPTABILITY
-
+

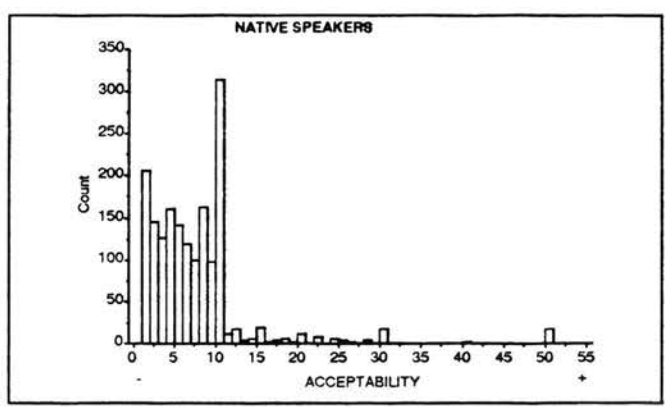
(d)


ENGLISH NEAR-NATIVE SPEAKERS
Count
ACCEPTABILITY
-
+

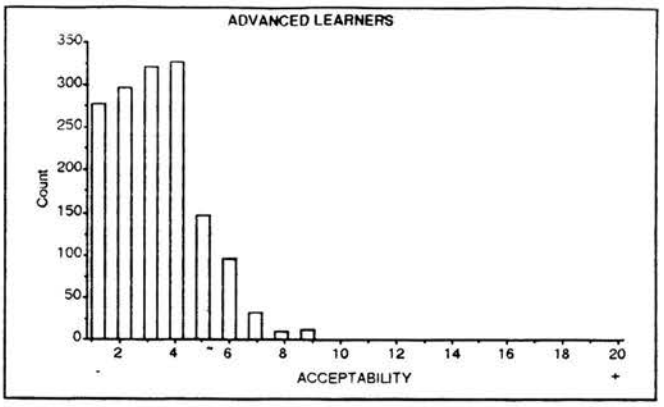
(e)


FRENCH NEAR-NATIVE SPEAKERS
Count
ACCEPTABILITY
-
+

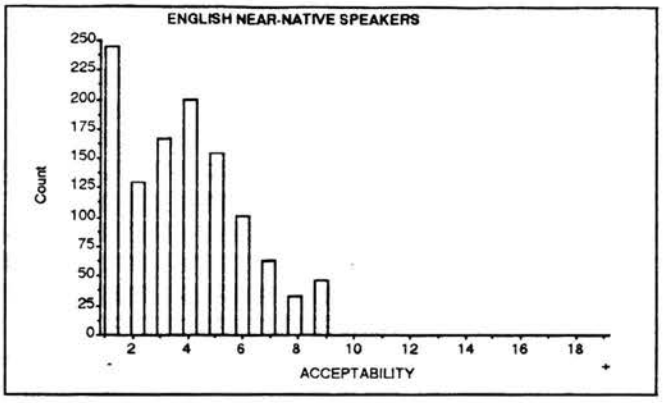
(f)


NATIVE SPEAKERS
Count
ACCEPTABILITY
-
+

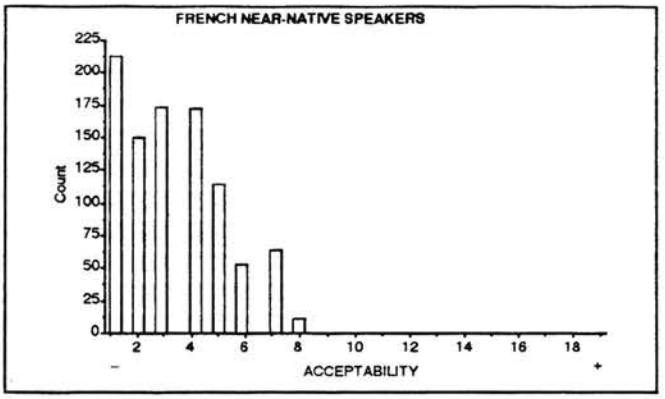
(c)


ADVANCED LEARNERS
Count
350
300
250
200
150
100
50
0
2
4
6
8
10
12
14
16
18
20
ACCEPTABILITY
-
+

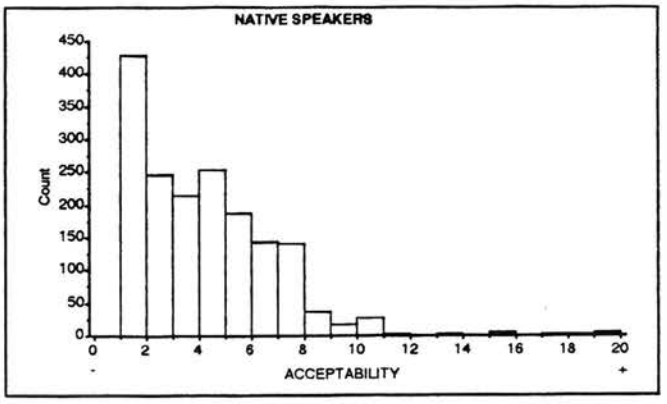
(d)


ENGLISH NEAR-NATIVE SPEAKERS
Count
250
225
200
175
150
125
100
75
50
25
0
2
4
6
8
10
12
14
16
18
ACCEPTABILITY
-
+

(e)


FRENCH NEAR-NATIVE SPEAKERS
Count
225
200
175
150
125
100
75
50
25
0
2
4
6
8
10
12
14
16
18
ACCEPTABILITY
-
+

(f)


NATIVE SPEAKERS
Count
450
400
350
300
250
200
150
100
50
0
0
2
4
6
8
10
12
14
16
18
20
ACCEPTABILITY
-
+

Although the Card-Sorting scores are spread over a narrower range than those from Magnitude Estimation, they present similar patterns. Non-native informants - regardless of proficiency level - use approximately the same range of scores (i.e. made the same number of piles). Native speakers use a wider range, although the majority of their scores are still concentrated from 1 to 4. What is different across levels is once again the increase in the percentage of scores at the lowest extreme of the range: thus, in the judgments of near-natives and natives 1 is the most frequently used score, whereas in the learners' judgments the mode is 4. The Card-Sorting data are therefore consistent with the Magnitude Estimation data in showing an increase across proficiency levels in the number of sentences judged as having low acceptability, although this effect is not as marked as with Magnitude Estimation .

7.2.2.3 Summary of findings

Two factors emerge from the frequency distributions of scores obtained with the two methods:

(a) subjects use a wider range of numerical values to represent relative acceptability with Magnitude Estimation than with Card-Sorting;

(b) with both methods, higher language proficiency implies more frequent use of low scores, and thus improvement in the recognition of sentence unacceptability.

7.3 Significant effects in the Analyses of Variance

Although both the results of Magnitude Estimation and those of Card-Sorting were subjected to the same statistical treatment (i.e. ANOVAs followed by Tukey tests of significance), only Magnitude Estimation produces an interval scale of measurement that legitimately allows the application of parametric statistics. Card-Sorting results in an ordinal scale, thus denying, in principle, access to this type of statistical analyses. However, in order to compare the two methods directly, it was decided to treat the Card-Sorting data as if they were on an interval scale. Relying on the identical research design underlying the collection of data with the two methods, two sets of parallel analyses were therefore conducted , to provide a common basis for comparison.

As a measure of the sensitivity of each method to variation in acceptability, we counted the number of statistically significant effects and interactions produced by Magnitude Estimation and Card-Sorting in the ANOVAs. Table 7.3 summarizes the total numbers of significant effects with their respective probability levels: it reports the totals of significant effects in the by-subject analysis (F_1 values), in the by-materials analysis (F_2 values) and, when both F_1 and F_2 are significant, in their combination (Min F'). The reader can refer to the complete ANOVA tables for Magnitude Estimation and Card-Sorting in Appendix B.

TABLE 7.3. Number of significant effects and interactions for Magnitude Estimation (ME) and Card-Sorting (CS) (all ANOVAs combined)

	F_1					F_2					MinF'		
	ns	*	**	***	****	ns	*	**	***	****	ns	*	**
ME	47	7	19	6	52	70	6	8	10	37	3	14	39
CS	56	2	13	11	48	75	6	8	10	32	12	11	25

**** = $p \leq .0001$; *** = $p \leq .001$; ** = $p \leq .01$; * = $p \leq .05$; ns = non significant

The overall numbers of significant effects obtained with Magnitude Estimation and Card-Sorting for each subset of linguistic materials judged by the informants (i.e. unergative, unaccusative, and restructuring verbs) , and for each grouping of subjects (i.e. all levels vs native/near-native), are reported in Tables 7.4, 7.5, 7.6.

TABLE 7.4. Total numbers of significant effects with **unergative** verbs for Magnitude Estimation (ME) and Card-Sorting (CS)

	ME			CS		
	F_1	F_2	Min F'	F_1	F_2	Min F'
Unergative, all variables, all levels	13	10	10	11	9	7
Unergative, all variables, native vs near-native	10	7	7	7	9	4
Unergative, auxiliary preferences, all levels	7	5	5	6	5	5
Unergative, auxiliary preferences, native vs near-native	4	3	3	4	3	2

TABLE 7.5. Total numbers of significant effects with **unaccusative** verbs for Magnitude Estimation (ME) and Card-Sorting (CS)

	ME			CS		
	F_1	F_2	Min F'	F_1	F_2	Min F'
Unaccusative, all variables, all levels	11	6	5	7	4	3
Unaccusative, all variables, native vs near-native	7	7	5	8	6	4
Unaccusative, auxiliary preferences, all levels	6	3	3	4	2	2
Unaccusative, auxiliary preferences, native vs near-native	3	2	2	2	2	2

TABLE 7.6. Total numbers of significant effects with **Restructuring** verbs for Magnitude Estimation (ME) and Card-Sorting (CS)

	ME			CS		
	F_1	F_2	Min F'	F_1	F_2	Min F'
Restructuring, all variables, all levels	6	7	4	9	6	4
Restructuring, all variables, native vs near-native	10	7	6	7	6	2
Restructuring, auxiliary preferences, all levels	4	2	2	6	3	1
Restructuring, auxiliary preferences, native vs near-native	3	2	2	3	1	0

The summary offered by Table 7.3, as well as inspection of Tables 7.4 - 7.6, indicate that the two methods do not produce inconsistent results. In general, if an effect or interaction is significant with Card-Sorting, it is also significant with Magnitude Estimation. However, there are significant effects and interactions with Magnitude Estimation that are not replicated with Card-Sorting. The reverse does not normally hold: only a small handful of effects turn out to be significant with Card-Sorting but not with Magnitude Estimation.

There are three ways in which Card-Sorting fails to yield the same pattern of significance as Magnitude Estimation:

(a) effects that are significant in their F_1 (or F_2) value with Magnitude Estimation do not reach significance in the corresponding analyses with Card-Sorting;

(b) effects that produce both a significant F_1 and a significant F_2 with Magnitude Estimation give only a significant F_1 (or F_2) with Card-Sorting;

(c) effects that produce a significant F_1 and F_2 with both methods give a significant MinF' with Magnitude Estimation but not with Card-Sorting.

The majority of significant effects found only by Card-Sorting are almost all concentrated in the judgments on Restructuring verbs, specifically the effects involving the Raising vs Control distinction: see Chapters 8 and 9 for a linguistic interpretation of why these particular

effects are not significant with Magnitude Estimation). Most of them are significant only with respect to F_1 but do not reach significance for F_2 or MinF'.

In order to determine whether the frequency patterns for significant effects found by the two methods differ significantly, Chi-square analyses were applied on the F_1 , F_2, and Min F' subsets of Table 7.3. The results are the following:

- total number of F_1 significant effects: Chi-square (4 d.f.) = 2.62, ns
- total number of F_2 significant effects: Chi-square (4 d.f.) = 2.42, ns
- total number of Min F' significant effects: Chi-square (2 d.f.) = 8.27, $p < .03$

These results again indicate the Magnitude Estimation and Card-Sorting produce broadly similar patterns of results: in fact there is no significant difference between the total numbers of F_1 and F_2 effects. Magnitude Estimation, however, is more sensitive to variation in linguistic acceptability , as shown by the significant difference between the total numbers of Min F' values that reached significance with the two methods.
This is not particularly surprising, given the fact that only the Magnitude Estimation data justify the application of inferential statistics: in running analyses of variance on the Card-Sorting data, we are pretending that this method produces an interval scale of measurement. We know that Card-Sorting is inherently less powerful than Magnitude Estimation because it cannot provide the information conveyed by an interval scale. What the comparison between the two methods has shown is that the additional information provided by Magnitude Estimation is meaningful, and not just noise (which is what Magnitude Estimation would have given if the informants had been as incapable of providing ratio judgments as they felt themselves to be).

7.4 Magnitude Estimation vs Card-Sorting: summary

The results that have been examined in this chapter suggest three general conclusions. First, linguistic acceptability appears to be a gradable dimension: informants clearly assigned different numbers to different sentence types, and distributed different sentence types into separate piles. This means that they perceived acceptability as a relative, rather than absolute, property of sentences.

Second, Magnitude Estimation turned out to be applicable to the elicitation of acceptability judgments from both native and non-native informants, despite its low face-validity. Its

administration did not present particular problems, was fast, convenient and not demanding in terms of technical equipment.

Third, Magnitude Estimation is at least as adequate as Card-Sorting as a method for the investigation of linguistic acceptability, since the results of the two procedures follow the same general trends; it is more powerful than Card-Sorting because, being more sensitive to variation in linguistic acceptability, it yields more significant effects and interactions among the experimental variables, and therefore a more fine-grained picture of variation in acceptability. Obviously, the higher sensitivity of Magnitude Estimation with respect to other methods is an advantage only if it produces an interpretable pattern of results . Chapters 8 and 9 describe the data obtained in this experiment in the framework of the theoretical hypotheses outlined in Chapter 4.

One of the interesting questions raised by this first application of Magnitude Estimation is why it did not elicit the wide range of scores that was theoretically expected. As we saw earlier, most informants consistently used a ten-point range. It is possible that informants had been conditioned to think of acceptability as 'seriousness of grammatical errors' to be marked on a ten-point scale (10 = very good, 1 = very bad). It may also be that linguistic acceptability (unlike metric continua) is a dimension of limited variation, along which only few degrees can be perceived. A definite answer to these questions will require more extensive validation of Magnitude Estimation as a procedure applied to acceptability judgments. But whatever the explanations for the narrowness of the informants' scale might be, they do not invalidate the results of this study because, as will be shown in Chapters 8 and 9, Magnitude Estimation not only produces a linguistically meaningful pattern of acceptability judgments, but a pattern that is in some cases more revealing than the one produced by Card-Sorting.

Chapters 8 and 9 therefore present only the results of Magnitude Estimation (and not of Card-Sorting) from the point of view of their linguistic significance in development and ultimate attainment. The corresponding results of Card-Sorting will be discussed only when they are relevant: the reader can otherwise assume that Card-Sorting is consistent with Magnitude Estimation. The full set of results of Card-Sorting can be found in Appendix C.

CHAPTER 8
RESULTS (2): DEVELOPMENTAL ASPECTS

8.0 Introduction

This chapter is concerned with the development of linguistic intuitions about auxiliary selection in Italian as a foreign language. It focuses on the acceptability judgments expressed by native Italian speakers, and on the judgments of English-speaking learners of Italian. The judgments of French near-native speakers are not taken into account here: they will be examined in Chapter 9.

8.1 Unergative verbs

Our analysis of results will start from the judgments obtained on sentences containing unergative verbs. Recall that, for this class of verbs, it was predicted that:

(a) because of the hypothesized lexical -semantic hierarchy subdividing the range of unergative verbs, auxiliary selection with non-motional verbs (i.e. *lavorare* ,'work') would be more determinate in informants' judgments than auxiliary selection with motional verbs (i.e. *camminare* ,'walk'), which in turn would be more determinate than motional verbs with an unaccusative alternant (i.e. *correre* ,'run'): this implies that the differences between auxiliaries in the informants' perception would be greatest with non-motional verbs and smallest with [+unaccusative alternant] verbs;

(b) there would be differences between the acceptability of unergative verbs in basic sentences and of the same verbs in Ne-cliticized sentences. Ne-cliticization with unergative verbs would be uniformly rejected as unacceptable, independently of verb category, because as a purely syntactic phenomenon it would not be sensitive to the lexical-semantic hierarchy;

(c) non-native judgments would approximate the values and pattern of native judgments with respect to auxiliary selection with unergatives in basic sentences but there would be no gradualness in the development of linguistic intuitions about unergatives in Ne-cliticized sentences, because the latter phenomenon should be more difficult to acquire.

In order to test these hypotheses, the data were subjected to the following statistical analyses:

(a) A pair of four-way ANOVAs with repeated measures (proficiency level x semantic category x auxiliary x word order), both by subjects and by materials;

(b) Min F' values were calculated for all pairs of significant F_1 and F_2 values;

(c) Post-hoc Tukey tests were carried out between pairs of F_1 means on the basis of the by subjects ANOVA results. The critical value of q at $p < .01$ for all comparison is 11.74.

In the following section (and in the rest of this chapter), F_2 and Min F' values will be reported only if significant. The full set of ANOVA tables (both by subjects and by materials) can be found in Appendix B.

8.1.1 All variables

The geometric means of the acceptability judgments provided by the four learner groups and the native speakers are reported in Table 8.1:

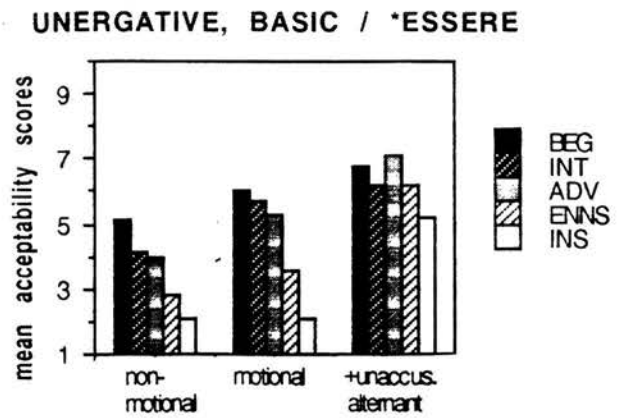
TABLE 8.1: Geometric means of acceptability judgments on unergative verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

	BEG	INT	ADV	ENNS	INS
NON-MOTIONAL, BASIC, ESSERE	5.133	4.106	4.007	2.810	2.067
NON-MOTIONAL, BASIC, AVERE	9.747	9.119	10.302	10.716	9.841
NON-MOTIONAL, NE-CL, ESSERE	6.621	4.653	5.334	2.779	2.054
NON-MOTIONAL, NE-CL, AVERE	6.751	6.689	7.302	7.943	3.105
MOTIONAL, BASIC, ESSERE	6.015	5.728	5.323	3.555	2.087
MOTIONAL, BASIC, AVERE	7.550	7.428	8.476	9.940	9.811
MOTIONAL, NE-CL, ESSERE	6.018	4.817	6.107	4.320	2.913
MOTIONAL, NE-CL, ESSERE	6.641	6.321	7.015	6.838	3.216
[+UNACCUSATIVE ALTERNANT], BASIC, ESSERE	6.810	6.176	7.074	6.241	5.206
[+UNACCUSATIVE ALTERNANT], BASIC, AVERE	7.471	7.201	7.616	6.420	7.542
[+UNACCUSATIVE ALTERNANT], NE-CL, ESSERE	5.813	5.516	7.297	6.329	5.515
[+UNACCUSATIVE ALTERNANT], NE-CL, AVERE	6.954	5.882	6.486	5.636	3.079

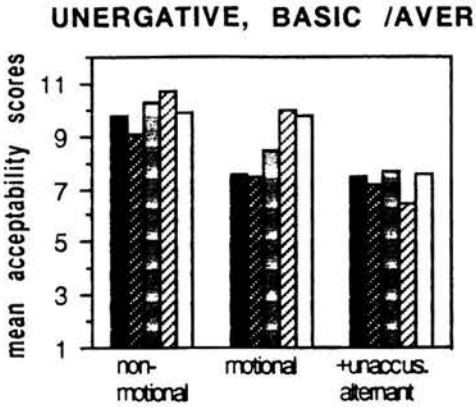
Fig. 8.1 offers a graphical representation of the data in the table: Figs. 8.1a, b represent judgments obtained for the sentences in the basic form, whereas Figs. 8.1c, d represent judgments on the sentences in the Ne-cliticization form:

FIGURE 8.1: Mean acceptability judgments on **unergative** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers).

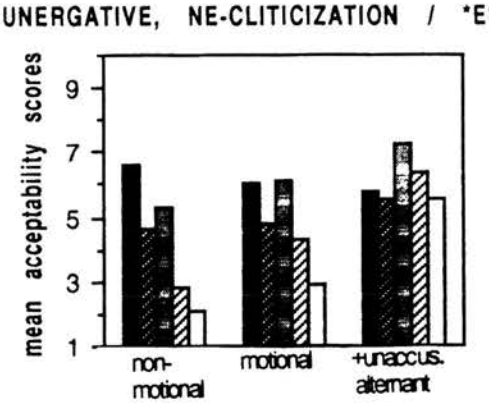
(a)



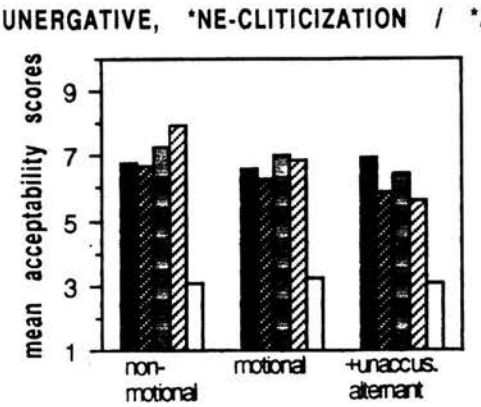
(b)



(c)



(d)



The graphs in Fig. 8.1 show three clear trends in the data. First, the five subject groups do not give uniform judgments: non-native acceptability values are generally different from native values, and there are further differences among the non-native judgments at various levels. Level of proficiency therefore is an important determinant of the particular acceptability pattern shown here. Statistically, this is confirmed by a repeated-measures ANOVA, which not only gives a main effect of proficiency level, $F_1(4, 155) = 12.74$, $p < .01$, $F_2 (4, 72) = 29.82$, $p < .0001$, Min F' $(4, 227) = 8.93$, $p < .01$, but also significant interactions between proficiency level and each of the three other variables: semantic category, $F_1 (8, 310) = 5.32$, $p < .01$, word order, $F_1 (4, 155) = 10.61$, $p < .0001$, $F_2 (4, 72) = 4.82$, $p < .002$, Min F' $(4, 139) = 3.31$, $p < .05$, auxiliary, $F_1 (4, 310) = 6.42$, $p < .0001$, $F_2 (4, 72) = 10.82$, $p < .0001$, Min F' $(4, 313) = 4.03$, $p < .01$. Further, there is a four-way interaction of semantic category, proficiency level, auxiliary and word order, $F_1 (8, 310) = 2.34$, $p < .02$. These results indicate that the five subject groups do not react uniformly on any of these variables.

Second, despite these effects, non-native judgments reproduce the same general pattern as native judgments (represented by the white bars in Fig. 8.1): the hypothesized differences among unergative verbs in terms of semantic categories are manifested by the intuitions of both native and non-native speakers. This is reflected in the repeated-measures ANOVA by a main effect of semantic category, $F_1 (2, 310) = 18.69$, $p < .01$, $F_2 (2, 18) = 6.84$, $p < .0001$, Min F' $(2, 33) = 5.01$, $p < .05$. The same analysis, however, also gives an interaction of semantic category by auxiliary, $F_1 (2, 310) = 70.89$, $p < .01$, $F_2 (2, 18) = 23.64$, $p < .0001$, Min F' $(2, 32) = 17.73$, $p < .01$, which suggests that the patterns of judgments for the two auxiliaries are not identical. As we shall see, the differences among semantic categories are reflected only by the judgments on sentences containing the ESSERE auxiliary. For the grammatical sentences with auxiliary AVERE, the three categories of unergatives are not perceived as different by the informants.

Third, the same pattern as the basic sentences is shown by the Ne-cliticization sentences (see Fig.8.1c and d). This runs counter the prediction that assumed Ne-cliticization to be equally unacceptable with either ESSERE or AVERE. However, the similarity between basic and Ne-cliticization word orders holds only with respect to ESSERE. This is partly due to the fact that, for Ne-cliticization sentences with 'non-motional' verbs, native speakers prefer AVERE, whereas they tend to accept ESSERE with Ne-cliticization sentences containing [+unaccusative alternant] unergatives. The ANOVA gives a significant main effect of word order, $F_1 (1, 155) = 44.28$, $p < .01$, $F_2 (1, 18) = 23.58$, $p < .0001$, Min F' $(1, 41) = 15.39$, $p < .01$, and no interaction between semantic category and word order. However, there is a significant interaction between word order and auxiliary, $F_1 (1, 155) = 81.63$, $p < .01$, $F_2 (1,$

18) = 27.54, p < .0001, Min F' (1, 32) = 20.51, p <.01, suggesting that the distinction between basic and Ne-cliticization sentences in the informants' judgments is a function of auxiliary: the two word orders are judged similarly when they include the ESSERE auxiliary and differently when they include the AVERE auxiliary.

Post-hoc Tukey tests show that native speakers do not distinguish between basic and Ne-cliticization versions of the same sentence with ESSERE (they are, predictably, treated as equally unacceptable). They do, however, distinguish between basic and Ne-cliticization sentences in their AVERE version, since AVERE is grammatical with basic sentences but ungrammatical with Ne-cliticization sentences: the BASIC vs NE-CL comparisons are significant for all the three semantic categories: non-motional (q = 36.12), motional (q = 34.93), and [+unaccusative alternant] (q = 28.05). No other significant difference results for any of the non-native subject groups.

The shape of the acceptability patterns found so far crucially depends on the auxiliary. This is reflected by a highly significant main effect of auxiliary, F_1 (1, 155) = 165.78, p < .01, F_2 (1, 18) = 86.46, p< .0001, Min F' (1, 40) = 56.82, p <.01.
Taking the ESSERE/AVERE comparisons as an indication of the informants' sensitivity to the appropriateness of auxiliary choice, one can easily see that the number of significant differences is a function of both (a) the level of proficiency, and (b) the type of semantic category of unergative verb.
At the beginning and intermediate levels (represented by the dark grey bars in Fig. 81), the only significant difference (q = 20.08 and q = 24.99 respectively) is for non-motional verbs in basic word order: ESSERE is judged more acceptable than AVERE. At the advanced level (represented by the light grey bars in Fig. 8.1) there are two significant differences: one for non-motional unergatives in basic sentences (q = 29.6) and the other for motional unergatives in basic sentences (q = 14.57). At the near-native level (represented by the white striped bars in Fig. 8.1), there are significant differences between the higher acceptability scores given to ESSERE and the lower scores given to AVERE for non-motional verbs both in basic form (q = 41.91) and in Ne-cliticization form (q = 32.87), as well as for motional verbs, again both in basic (q = 32.19) and in Ne-cliticization form (q = 14.38).
Only at the native level (represented by the white bars in Fig. 8.1) does one find significant differences for all the three semantic categories in both word orders: non-motional (basic: q = 48.86; Ne-cliticization: q = 12.94), motional (basic: q = 48.46; Ne-cliticization: q = 11.93), and [+unaccusative alternant] (basic: q = 11.88; Ne-cliticization: q = 18.25).

Thus, it appears that sensitivity to auxiliary choice with unergative verbs first manifests itself in the basic word order, and only later extends to Ne-cliticization contexts: learners have to acquire the knowledge that Ne-cliticization is disallowed with unergatives. Surprisingly, it is only the native speakers who are able to differentiate between correct use of AVERE in basic sentences and incorrect use of AVERE in Ne-cliticization sentences.

The fourth clear tendency shown by Fig. 8.1 is the gradual development of non-native acceptability values from least to most experienced learners. As evidence for this, one can turn to the Tukey tests on inter-category comparisons: these show that the overall number of significant differences among types of unergative verbs increases with knowledge of Italian. Moreover, the first differences to be found at lower proficiency levels are those between most and least typical verb category, and those concerning basic sentences.

For beginners, there are no significant differences among categories: this group of learners do not differentiate within the class of unergative verbs. At the intermediate and advanced levels, the only significant difference is between non-motional and [+unaccusative alternant] unergative verbs in basic sentences with basic word order ($q = 12.78$ and $q = 17.8$, respectively), where the former are given lower acceptability ratings than the latter.

As knowledge of Italian improves, learners begin to discriminate between most and least 'typical' category of unergative verbs. At the near-native level, there are significant differences between non-motional and [+unaccusative alternant] unergatives with the ESSERE auxiliary in both basic ($q = 24.98$) and Ne-cliticization form ($q = 16.49$). There is also a significant difference between the same two semantic types of unergatives with the AVERE auxiliary, but only in basic sentences ($q = 11.75$). Near-natives also distinguish in a statistically significant way between motional and [+unaccusative alternant] verbs with ESSERE in both basic ($q = 17.61$) and Ne-cliticization sentences ($q = 11.96$), and with AVERE in basic sentences only ($q = 13.69$).

The judgments of native Italian speakers (represented by the white bars in Fig.8.1a) make a clear distinction between non-motional and motional unergatives, on the one hand, and '[+unaccusative alternant]' verbs on the other. This is evident with respect to sentences in their basic form including the ESSERE auxiliary : unpaired unergative verbs elicit acceptability judgments that are not different from each other, but are different from the judgments produced on paired unergative verbs, which are perceived as more acceptable. Accordingly, the Tukey tests indicate significant differences between non-motional and [+unaccusative alternant] verbs with ESSERE as an auxiliary, in both basic ($q = 28.93$) and

Ne-cliticization sentences (q = 30.93), and between motional and [+unaccusative alternant] unergatives, again in both basic (q = 28.62) and Ne-cliticization sentences (q = 19.98): in both cases, [+unaccusative alternant] verbs are given significantly higher acceptability ratings when they are presented with ESSERE than the other two verb types.

A more direct measure of the gradual progress of non-native judgments towards the values of native ones is given by the comparisons among proficiency level.
Predictably, the strongest inter-group differences are those between highest and lowest proficiency levels, and particularly those that separate learners from near-native/native speakers. Furthermore, the greatest number of significant differences are to be found with respect to non-motional unergatives; the lowest number with respect to [+unaccusative alternant] unergatives. For non-motional verbs in basic sentences with the incorrect auxiliary ESSERE (see Fig. 8.1a), beginners' judgments are significantly higher than near-natives' (q = 18.86) and natives' (q = 28.48). The judgments of learners at the intermediate level also differ , in the same direction, from those of near-natives (q = 11.87) and from those of native speakers (q = 21.49). Finally, the judgments of advanced learners are still different from those of native speakers (q = 20.73). The same pattern of differences is obtained for non-motional verbs selecting ESSERE in Ne-cliticization sentences: near-natives offer lower acceptability ratings than beginners (q = 27.18), intermediate learners (q = 16.13) and advanced learners
(q = 20.41); native speakers' judgments are significantly lower than beginners (q =36.66), intermediate learners (q = 25.61) and advanced learners (q = 29.89).

All the non-native groups differ from native speakers on all the three semantic types of unergatives in Ne-cliticization sentences with AVERE, as can be seen in Fig. 8.1d: non-natives tend to accept these sentences, whereas native speakers reject them. This once again suggests that it takes time for learners to acquire the ungrammaticality of Ne-cliticization with unergative verbs.
Native speakers differ from beginners (non-motional: q = 24.32; motional: q = 22.71; [+unaccusative alternant]: q = 25.51), from learners at the intermediate level (non-motional: q = 24.03; motional: q = 21.16; [+unaccusative alternant]: q = 20.27), from advanced learners (non-motional: q = 26.78; motional: q = 24.42; [+unaccusative alternant]: q = 23.33) and from near-native speakers (non-motional: q = 29.41; motional: q = 23.62; [+unaccusative alternant]: q = 18.93).

As for Ne-cliticized sentences with ESSERE (see Fig. 8.1c), Tukey tests again separate non-natives, who tend to accept them, from native speakers,who tend to reject them. Thus,

native speakers give significantly lower acceptability ratings to these sentences with motional unergatives than beginners (basic: q = 33.1; Ne-clticization: q = 22.71), intermediate (basic: q = 31.61; Ne-cliticization: q = 31.61), advanced learners (basic: q =29.31; Ne-cliticization: q = 23.17), and near-native speakers (basic: q = 29.31; Ne-cliticization: q = 12.33). There are also differences between the low acceptability ratings of near-natives and the higher ones of beginners (basic: q = 16.46; Ne-cliticization: q = 23.84), intermediate learners (basic: q =14.93; Ne-cliticization: q = 14.93), and advanced learners (basic: q = 12.63).

8.1.2 Auxiliary preferences

To assess sensitivity to the AVERE-ESSERE distinction directly, an ANOVA was conducted on the mean differences between judgments on the correct AVERE auxiliary and those on the incorrect ESSERE auxiliary. The prediction was that non-motional unergative verbs would be associated with stronger auxiliary preferences than those for motional unergative verbs, and that [+unaccusative alternant] unergatives would be associated with the weakest auxiliary preferences. This prediction is motivated by the hypothesis that non-motional unergatives (which denote activities with no implicit change-of-location element) are the core type of unergative, whereas [+unaccusative alternant] verbs are the most peripheral type.

In order to test these predictions, the following statistical analyses were carried out on the mean differences between auxiliaries:

(a) a pair of three-way repeated-measures ANOVAs (proficiency level x semantic category x word order);
(b) Min F' values were calculated for each pair of significant F_1 and F_2 values;
(c) Tukey tests of significance between F_1 means were applied on the basis of the ANOVA results. The critical value of q at $p < .01$ is 18.69.

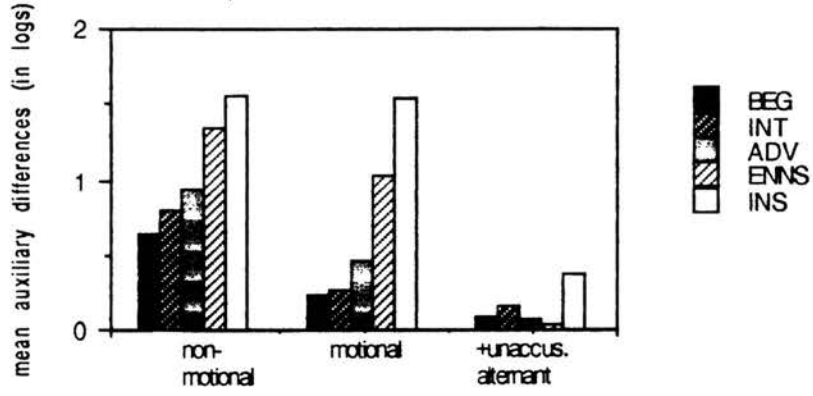
The results are reported in Table 8.2 and their graphic representation is displayed in Figs 8.2 a, b. The data in Table 8.2 are reported in logarithmic form because of the impossibility of exponentiating negative numbers. Positive numbers indicate that AVERE was found more acceptable than ESSERE; negative numbers indicate that ESSERE was preferred to AVERE. Larger numbers, positive or negative, indicate stronger preferences.

TABLE 8.2: Unergative verbs, all levels: mean auxiliary differences (AVERE - ESSERE) (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

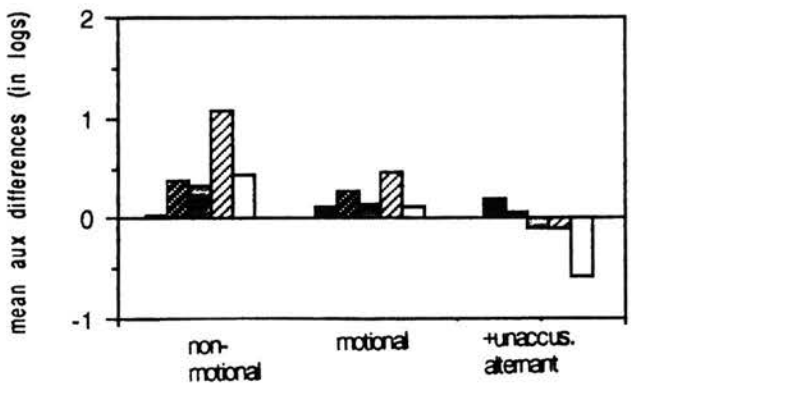
	BEG	INT	ADV	ENNS	INS
NON-MOTIONAL, BASIC	0.641	0.798	0.944	1.339	1.561
NON-MOTIONAL, NE-CLITICIZATION	0.019	0.363	0.314	1.050	0.413
MOTIONAL, BASIC	0.227	0.260	0.465	1.028	1.548
MOTIONAL, NE-CLITICIZATION	0.099	0.272	0.139	0.459	0.099
[+UNACCUSATIVE ALTERNANT], BASIC	0.093	0.154	0.074	0.028	0.371
[+UNACCUSATIVE ALTERNANT], NE-CL	0.179	0.064	-0.118	-0.116	-0.583

FIGURE 8.2: Unergative verbs, all levels: mean auxiliary preferences (AVERE-ESSERE) for (a) basic and (b) Ne-cliticization sentences (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

(a)



(b)



Figs 8.2a, b effectively summarize all the trends revealed by the previous analysis. The graphs suggest that the ability to differentiate among different types of unergative verbs develops with proficiency. This explains why the significant main effects obtained in the ANOVA for both proficiency level, F_1 (4,155) = 6.42, $p < .0001$, F_2 (4, 72) = 10.82, $p < .0001$, Min F' (4, 223) = 4.03, $p < .01$, and semantic category, F_1 (2, 310) = 70.89, $p < .0001$, F_2 (2,18) = 23.64, $p < .0001$, Min F' (2, 32) = 17.33, $p < .01$, are qualified by a significant interaction of semantic category and proficiency level, F_1 (8, 310) = 6.23, $p < .0001$, F_2 (8, 72) = 6.44, $p < .0001$, Min F' (8, 239) = 3.17, $p < .01$, and by a three-way interaction of semantic category, word order and proficiency level, F_1 (8, 310) = 2.34, $p < .02$.

Tukey tests on the comparisons among categories provide further evidence for the gradual development of non-native intuitions. There is no significant difference among auxiliary preferences for different verb categories at the beginning level. At the intermediate and advanced levels, preferences for AVERE are weaker with [+unaccusative alternant] verbs than with non-motional verbs in basic sentences ($q = 20.14$ and $q = 27.21$, respectively). At the near-native and native levels, AVERE is judged as significantly less acceptable (only in basic sentences) with [+unaccusative alternant] verbs than with non-motional verbs (near-native: $q = 40.94$; native: $q = 37.19$) and with motional verbs, (near-native: $q = 31.24$; native: $q = 36.79$). There are no significant preferences among unergative verb types in Ne-cliticization sentences.

As Figs 8.2a and b show, the strongest auxiliary preferences are in the area of non-motional unergatives, followed by motional unergatives and last by [+unaccusative alternant] verbs. The judgments of all groups reflect this pattern for the three types of unergative verbs. Irrespective of proficiency level, all subjects perceive non-motional verbs as the most clearly acceptable with auxiliary AVERE and the most clearly unacceptable with auxiliary ESSERE. This is particularly the case at lower proficiency level, where there is a noticeable gap between the auxiliary difference for non-motional verbs and that for motional verbs. The gap is reduced with proficiency level and virtually disappears in the judgments of native speakers. The auxiliary difference for [+unaccusative alternant] verbs remains small across proficiency levels: subjects judge auxiliary AVERE with these verbs as only marginally more acceptable than ESSERE.

All groups assign higher acceptability scores to the AVERE auxiliary in basic sentences (as shown by the fact that all differences are positive), but in Ne-cliticization sentences there is a preference for ESSERE in more proficient subject groups, despite the theoretical ungrammaticality of both auxiliaries. Thus, for basic sentences there is a gradual change of

non-native judgments in the direction of native judgments. The pattern is less clear for the Ne-cliticization sentences, where only the most advanced proficiency levels show a decrease in the acceptability of AVERE across semantic categories. This is statistically supported in the ANOVA by an interaction of semantic category and word order, F_1 (2, 310) = 9.08, $p < .0001$, and by small but significant interaction of semantic category, word order and proficiency level, $F_1(8, 310) = 2.34$, $p < .02$.

These findings clearly point to a contrast between the pattern of judgments for basic sentences and that for Ne-cliticization sentences, as confirmed in the ANOVA by a significant main effect of word order, F_1 (1, 155 = 81.63, $p < .0001$, $F_2(1, 18) = 27.54$, $p < .0001$, Min F' (1, 32) = 20.59, $p < .01$. This difference, however, is not uniformly present in native and non-native judgments, as suggested by a strong interaction of word order and proficiency level, F_1 (4, 155) = 15.01, $p < .0001$, F_2 (4, 72) = 13.85, $p < .0001$, Min F' (4, 190) = 7.2, $p < .01$.

Post-hoc Tukey tests indicate the reason: only native speakers can differentiate between basic and Ne-cliticization sentences containing unergative verbs, as we saw earlier. These tests also suggest that the starting point for the development of this ability is the non-motional verbs. The BASIC vs NE-CLITICIZATION comparisons are significant only for non-motional unergative verbs at the lowest proficiency level (q = 19.43) and at the advanced level (q = 19.7). None of the comparisons is significant at the intermediate level. At the near-native level, the significant differences are both for non-motional verbs (q = 19.02) and for motional verbs (q = 18.78). Only native speakers produce significantly different preferences across word order for all types of verbs: non-motional (q = 35.85), motional (q = 45.28), and [+unaccusative alternant] (q = 29.8).

The between-groups comparisons again show a clearcut division between learners on the one hand (including beginner, intermediate and advanced levels), and near-native/native speakers on the other. For basic sentences, the same pattern of differences can be observed with respect to both non-motional and motional unergatives. Near-natives have larger auxiliary differences (and therefore a stronger preference for AVERE) than beginners (non-motional: q = 21.79; motional: q = 25.02), and intermediate learners (non-motional: q = 16.89; motional: q = 24.01), whereas native judgments indicate larger auxiliary differences than beginners (non-motional: q = 28.73; motional: q = 42.26), intermediate learners (non-motional: q = 23.83; motional: q = 40.25), and advanced learners (non-motional: q = 19.26; motional: q= 33.83). No significant differences between groups were found with respect to [+unaccusative alternant] unergatives in basic sentences.

There are fewer significant differences for Ne-cliticization sentences. None of them concerns motional unergative verbs. Most of them are found for [+unaccusative alternant] verbs: all non-native groups have significantly larger auxiliary differences for this verb category than native speakers (beginners: q = 23.82; intermediate: q = 20.23; advanced: q = 14.54; near-native: q = 14.6). With non-motional verbs, the auxiliary differences obtained for near-natives are significantly larger than those of both beginners (q = 32.21) and intermediate learners (q = 21.47). This reflects the relatively higher acceptability scores given by near-natives to non-motional verbs with auxiliary AVERE in Ne-cliticization word-order.

8.1.3 Summary of findings

The results obtained for unergative verbs confirm the prediction that this class of verbs is not homogeneous. The hypothesized distinctions among non-motional, motional, and [+unaccusative alternant] verbs are reflected by the informants' judgments on auxiliary selection in basic sentences: non-motional verbs appear to be the most prototypically unergative, since they are judged as least acceptable with the auxiliary ESSERE. Conversely, [+unaccusative alternant] verbs are the least prototypically unergative, for they receive significantly higher acceptability ratings when they are presented with ESSERE. In terms of acquisition of auxiliary selection, non-motional verbs are the first ones to be acquired, followed by motional verbs, and last by +unaccusative verbs. The development of knowledge of auxiliary selection is thus conditioned by the semantic hierarchy we hypothesized in Chapter 3.

Contrary to the predictions, Ne-cliticization also seems to be influenced by semantic distinctions both in native and in non-native judgments. Its degree of acceptability with unergative verbs with ESSERE is itself a function of semantic prototypicality: [+unaccusative alternant] verbs are judged as significantly more acceptable in Ne-cliticization sentences than the other two verb categories. This suggests that syntactic reflexes of unaccusativity , far from being impervious to semantic characterizations, are strongly interconnected with them.

8.2 Unaccusative verbs

Let us now turn to the analysis of judgments given on unaccusative verbs. It will be recalled that, for this class of verbs, it was predicted that:

(a) the degree of determinacy of native judgments on auxiliary selection would correspond to the following hierarchy (where '>' means 'more determinate than'):

change-of-location > continuation-of-state > existence-of-state > [+transitive alternant] >[+unergative alternant]

The differences between auxiliaries in the perception of native speakers would be greatest for change-of-location verbs (i.e. *venire*, 'come') and smallest for [+unergative alternant] verbs (i.e. *correre (a casa)*, 'run (home').

(b) native speakers would accept both basic and Ne-cliticized sentences containing unaccusative verbs, but Ne-cliticized sentences - unlike basic ones - would be uniformly judged as acceptable, irrespective of verb category, because Ne-cliticization should not be sensitive to lexical-semantic representations.

(c) non-native judgments would approximate the values and pattern of native judgments for basic sentences, but there would be no gradual improvement with respect to Ne-cliticized sentences, because phenomena which do not have a lexical-semantic basis - such as NE-cliticization - should be more difficult to acquire than phenomena which do.

These hypotheses were investigated by means of the following statistical analyses:

(a) a pair of four-way repeated-measures ANOVAs by subject (proficiency level x semantic category x word order x auxiliary);

(b) Min F' values were computed for all pairs of significant F_1 and F_2 values;

(c) Tukey tests on post-hoc comparisons between the F_1 means. The critical value of q at $p < .01$ for all comparisons is 6.96.

8.2.1 All variables

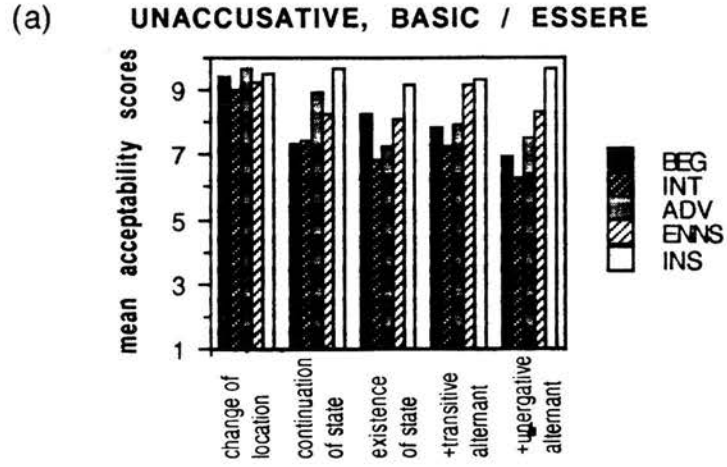
The geometric means of the acceptability judgments given by subjects on the five categories of unaccusative verbs are reported in Table 8.3.

TABLE 8.3: Unaccusative verbs, all levels, mean acceptability judgments (BEG = beginner; INT = intermediate; ADV = advanced; ENNS = English near-native, INS = Italian native)

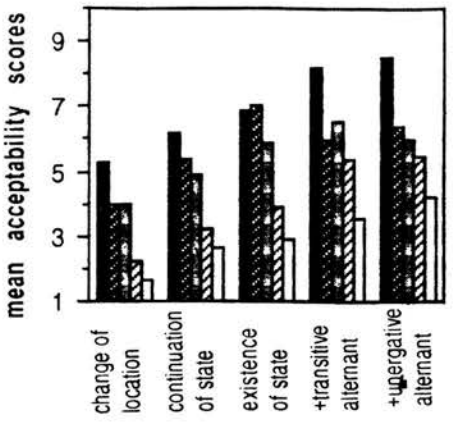
	BEG	INT	ADV	ENNS	INS
CHANGE-OF-LOCATION, BASIC, ESSERE	9.392	9.02	9.644	9.248	9.509
CHANGE-OF-LOCATION, BASIC, AVERE	5.316	3.973	3.983	2.224	1.653
CHANGE-OF-LOCATION, NE-CL, ESSERE	6.973	7.161	8.787	9.323	8.54
CHANGE-OF-LOCATION, NE-CL, AVERE	3.199	2.838	2.763	2.038	1.697
CONTINUATION-OF-STATE, BASIC, ESSERE	7.328	7.453	8.921	8.228	9.704
CONTINUATION-OF-STATE, BASIC, AVERE	6.188	5.364	4.847	3.214	2.665
CONTINUATION-OF-STATE, NE-CL, ESSERE	7.505	6.78	8.201	8.332	8.465
CONTINUATION-OF-STATE, NE-CL, AVERE	3.143	3.732	3.466	3.017	1.975
EXISTENCE-OF-STATE, BASIC, ESSERE	8.277	6.866	7.265	8.088	9.203
EXISTENCE-OF-STATE, BASIC, AVERE	6.877	7.065	5.892	3.93	2.922
EXISTENCE-OF-STATE, NE-CL, ESSERE	7.024	6.444	7.626	6.874	7.198
EXISTENCE-OF-STATE, NE-CL, AVERE	4.968	4.494	4.759	3.914	2.667
[+TRANSITIVE ALTERNANT], BASIC, ESSERE	7.851	7.284	7.898	9.17	9.34
[+TRANSITIVE ALTERNANT], BASIC, AVERE	8.217	5.954	6.553	5.353	3.562
[+TRANSITIVE ALTERNANT], NE-CL, ESSERE	6.766	5.4	6.575	8.013	7.241
[+TRANSITIVE ALTERNANT], NE-CL, AVERE	5.266	5.734	5.564	5.798	3.911
[+UNERGATIVE ALTERNANT], BASIC, ESSERE	6.934	6.299	7.543	8.367	9.686
[+UNERGATIVE ALTERNANT], BASIC, AVERE	8.551	6.392	5.94	5.448	4.204
[+UNERGATIVE ALTERNANT], NE-CL, ESSERE	5.411	5.396	6.505	8.331	8.08
[+UNERGATIVE ALTERNANT], NE-CL, AVERE	6.107	5.232	5.578	5.344	3.715

Fig. 8.3a-d give a visual representation of the data:

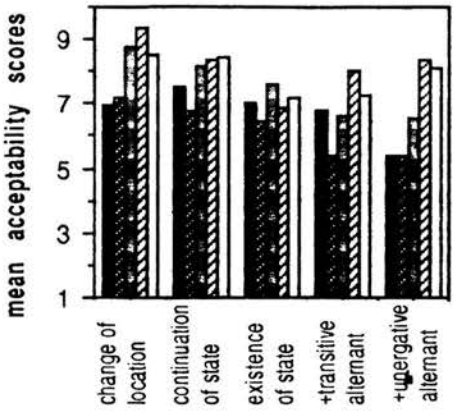
FIGURE 8.3: Mean acceptability judgments on unaccusative verbs (BEG = beginner; INT = intermediate; ADV = advanced; ENNS = English near-native, INS = Italian native)



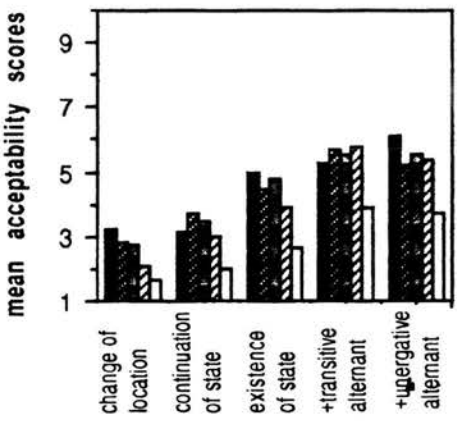
(b) UNACCUSATIVE, BASIC / *AVERE



(c) UNACCUSATIVE, NE-CLITICIZATION / ESSERE



(d) UNACCUSATIVE, NE-CLITICIZATION /*AVERE



The results suggest that unaccusative verbs are given different acceptability scores by informants, depending on their semantic type. This is evident from Fig. 8.3b and d, which represent the pattern of judgments on unaccusative verbs with the incorrect auxiliary AVERE in basic and Ne-cliticization sentences: change-of-location verbs are perceived as least

be judged in same way irrespective of verb category. On the whole, this progression is more marked for basic sentences than for Ne-cliticization sentences, as shownby a significant main effect of word order, F_1 (1, 155) = 67.34, $p < .0001$, F_2(1, 30) = 10.51, $p < .003$, Min F' (1, 40) = 9.12, $p < .01$, and by a significant interaction of word order, auxiliary and proficiency level, F_1 (4, 155) = 3.77, $p < .006$, F_2(4, 120) = 2.44, $p < .05$.

At the beginners' level (represented by the black bars in Fig. 8.3), there are no significant differences between auxiliaries for any of the verb categories: this group of learners clearly have very little knowledge of auxiliary selection. At the intermediate level (represented by the dark grey bars in Fig. 8.3), there are significant differences with respect to change-of-location verbs, both in basic (q = 11.73) and in Ne-cliticization sentences (q = 12.56), as well as with respect to continuation-of-state verbs in Ne-cliticization form (q = 8.11): in all these cases, ESSERE is judged more acceptable than AVERE.
At the advanced level (represented by the light grey bars in Fig. 8.3), there are significant differences with respect to both change-of-location and continuation-of-state verbs, both in basic (q = 12 and q = 8.28, respectively) and in Ne-cliticization word order (q = 15.71 and q = 11.69 respectively). At the near-native level (the white striped bars in Fig. 8.3), one finds significant differences for change-of-location verbs (basic: q = 19.34; Ne-cliticization: q = 20.64), continuation-of-state (basic: q = 12.76; Ne-cliticization: q = 13.78), existence-of-state (basic: q = 9.79; Ne-cliticization: q = 7.64), and [+transitive alternant] unaccusatives, in basic sentences only (q = 7.3). At the native level (represented by the white bars in Fig. 8.3), there are significant differences between high acceptability ratings for ESSERE and low acceptability ratings for AVERE for all verb categories in both word orders: change-of-location (basic: q = 23.75; Ne-cliticization: q = 21.93), continuation-of-state (basic: q = 17.54; Ne-cliticization: q = 19.75), existence-of-state (basic: q = 15.57; Ne-cliticization: q= 13.48), [+transitive alternant] (basic: q = 13.08; Ne-cliticization: q = 8.36), and [+unergative alternant] (basic: q = 11.33; Ne-cliticization: q = 10.55).

Despite the broad similarities between the pattern of judgments obtained for basic and Ne-cliticization sentences, knowledge of the relationship between unaccusativity and Ne-cliticization also improves with proficiency. Learners gradually acquire the knowledge that Ne-cliticization is possible with unaccusative verbs. The ANOVA in fact gives significant interactions of word order and proficiency level, F_1(4, 155) = 4.35, $p < .002$, word order and auxiliary, F_1 (4, 155) = 4.66, $p < .003$, and word order, auxiliary and proficiency level, F_1 (4, 155) = 3.77, $p < .006$, F_2 (4, 120) = 2.44, $p < .05$.
The pattern of significant differences given by post-hoc Tukey tests confirms the prediction that subjects at higher proficiency levels would not differentiate between the two

constructions, since they are both grammatical with auxiliary ESSERE and ungrammatical with auxiliary AVERE.
The type and distribution of inter-category differences is consistent with the overall picture we have outlined. In addition, they show that the ability to discriminate among categories of unaccusative verbs starts from Ne-cliticization sentences, and later is extended to basic sentences. It is possible that learners acquire knowledge of the intimate link between unaccusativity and Ne-cliticization early, so that Ne-cliticization sentences with AVERE are perceived as representing a double violation of ESSERE-selection.

As we mentioned before, all significant differences concern unaccusatives with AVERE. Few significant differences are found in the three groups of learners, and most of these concern Ne-cliticization sentences. It is only at the most advanced levels that informants are sensitive to the "centrality" of some verb types. So, at the beginners' level, there are differences in Ne-cliticization sentences with AVERE (see Fig. 8.3d) between change-of-location unaccusatives and [+unergative alternant] ($q = 8.78$), and between continuation-of-state verbs and both [+transitive alternant] verbs ($q = 7.01$) and [+unergative alternant] verbs ($q = 9.02$): verbs that occupy a higher position along the hierarchy are perceived as significantly less acceptable with AVERE than more peripheral verbs.

At the intermediate level, there is a difference between the lower acceptability ratings given to change-of-location verbs with AVERE and the higher ones given to existence-of-state verbs with AVERE, in basic word order ($q = 7.81$); there is also a difference, in the same direction, between change-of-location and [+transitive alternant] verbs in Ne-cliticization word order ($q = 9.54$). At the advanced level, informants judge change-of-location unaccusatives in Ne-cliticization sentences significantly less acceptable with AVERE than existence-of-state verbs ($q = 7.38$), [+transitive alternant] verbs ($q = 9.5$), and [+unergative alternant] verbs ($q = 9.54$).
Near-natives distinguish among categories of unaccusative verbs not only in Ne-cliticization sentences, but also in basic ones (see Fig. 8.3b). For basic word order, they judge change-of-location verbs with AVERE less acceptable than existence-of-state verbs ($q = 7.73$), [+transitive alternant] verbs ($q = 11.92$), and [+unergative alternant] verbs ($q = 12.16$) with the same auxiliary. They also judge continuation-of-state unaccusatives less acceptable than [+unergative alternant] verbs ($q = 7.17$). For Ne-cliticization sentences, there are again differences between change-of-location verbs and existence-of-state ones ($q = 8.86$), [+transitive alternant] ($q = 14.19$) and [+unergative alternant] ($q = 13.09$).
Finally, in the judgments of native speakers there are significant differences, for both basic and Ne-cliticization sentences, between the lower acceptability scores given to change-of-

location verbs and those given to existence-of-state verbs (basic: q = 7.73; Ne-cliticization: q = 7.31), [+transitive alternant] verbs (basic: q = 10.42; Ne-cliticization: q = 11.33), and [+unergative alternant] verbs (basic: q = 12.67; Ne-cliticization: q = 10.63). Furthemore, there are differences between the lower acceptability of continuation-of-state unaccusatives and that of both [+transitive alternant] verbs (q = 9.27) and [+unergative alternant] ones (q = 8.57). The pattern revealed by this analysis therefore confirms the link between knowledge of Italian and sensitivity to lexical semantic distinctions among unaccusative verbs.

The overall pattern of inter-group differences shows that all significant results are concentrated in the area of the three topmost categories of unaccusative verbs (i.e. change-of-location, continuation-of-state and existence-of-state): the five subject groups give similar responses on the two most peripheral categories (i.e. [+transitive alternant] and [+unergative alternant]). All significant differences concern responses on sentences with auxiliary AVERE.

For change-of-location verbs, natives and near-natives give lower acceptability ratings than learners. For basic sentences, near-native speakers offer significantly lower acceptability ratings than beginners (q = 11.83), intermediate learners (q = 7.87), and advanced learners (q = 7.91); similarly, native speakers' acceptability scores on this type of unaccusatives are lower than those of beginners (q = 15.86), intermediate learners (q = 11.9), and advanced learners (q = 11.94). Native judgments are also significantly lower than those of beginners (q = 8.6) and intermediates (q = 6.98) on Ne-cliticization sentences.

For continuation-of-state sentences in basic word order, near-natives give lower responses than beginners (q = 8.9) and intermediate learners (q = 6.96); native speakers' judgments are significantly lower than beginners (q = 11.44), intermediate learners (q = 9.5), and advanced learners (q = 8.12). Ne-cliticization sentences within the same category register differences between the lower acceptability ratings given by native speakers and those given by intermediate learners (q = 8.63), and advanced learners (q = 7.63).
Finally, existence-of-state unaccusative verbs in basic sentences are judged less acceptable with AVERE by near-natives than they are by beginners (q = 7.59) and intermediate learners (q = 7.95). There are significant differences between the lower judgments of native speakers and the higher judgments of beginners (q = 11.62), intermediates (q = 11.98), and advanced learners (q = 9.52). For Ne-cliticization sentences, native speakers' acceptability ratings are lower than beginners' (q = 8.45), intermediates' (q = 7.08), and advanced learners' (q = 7.86).

8.2.2 Auxiliary preferences

An ANOVA was run on mean differences obtained by subtracting the incorrect auxiliary AVERE from the correct ESSERE: as with unergative verbs, these differences provide a more direct indication of the sensitivity to auxiliary selection with unaccusative verbs. It was predicted that (a) the strength of auxiliary preferences in basic sentences would be proportional to the position of verb categories along the unaccusative hierarchy: they would be strongest for change-of-location verbs and weakest for [+unergative alternant] verbs; (b) the strength of auxiliary preferences for Ne-cliticization sentences would be uniform across verb categories, and (c) native auxiliary differences would be generally larger than non-native ones, and would maximally differ from beginners' and minimally differ from near-natives'.

These predictions were statistically analysed by means of :

(a) a pair of three-way repeated-measures ANOVAs (proficiency level x semantic category x word order);

(b) Min F' values, for each pair of significant F_1 and F_2 values;

(c) Post-hoc Tukey tests of significance applied on the F_1 means. The critical value of q for all comparisons is 8.93.

The mean differences are reported in Table 8.4 (in logarithmic form). Positive numbers indicate preference for ESSERE, negative numbers indicate preference for AVERE, and the size of numbers indicates the strength of preference.

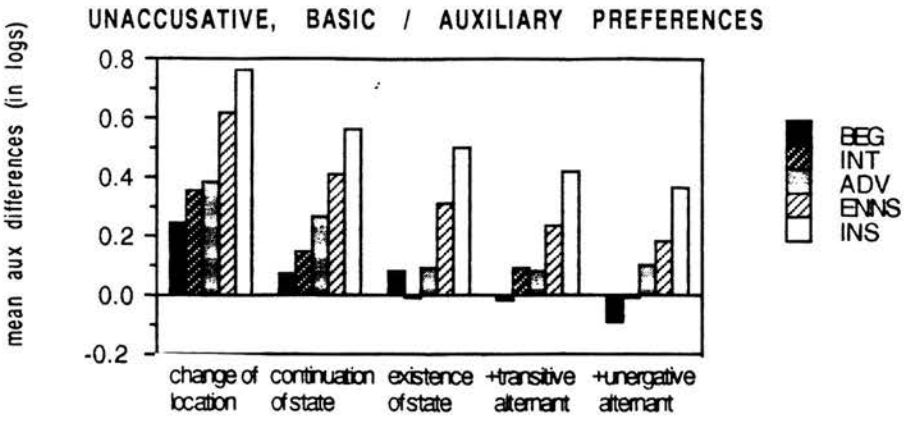
TABLE 8.4: Unaccusative verbs, mean auxiliary preferences (ESSERE - AVERE)
(BEG = beginner; INT = intermediate; ADV = advanced; ENNS = English near-native, INS = Italian native)

	BEG	INT	ADV	ENNS	INS
CHANGE-OF-LOCATION, BASIC	0.247	0.356	0.384	0.619	0.760
CHANGE-OF LOCATION, NE-CLITICIZATION	0.338	0.402	0.503	0.660	0.702
CONTINUATION-OF-STATE, BASIC	0.073	0.143	0.265	0.408	0.561
CONTINUATION-OF-STATE, NE-CLITICIZATION	0.378	0.259	0.374	0.441	0.632
EXISTENCE-OF-STATE, BASIC	0.080	-0.012	0.091	0.313	0.498
EXISTENCE-OF-STATE, NE-CLITICIZATION	0.150	0.157	0.205	0.245	0.431
[+TRANSITIVE ALTERNANT], BASIC	-0.020	0.088	0.081	0.234	0.419
[+TRANSITIVE ALTERNANT], NE-CLITICIZATION	0.109	-0.026	0.073	0.141	0.268
[+UNERGATIVE ALTERNANT], BASIC	-0.091	-0.006	0.104	0.186	0.362
[+UNERGATIVE ALTERNANT], NE-CLITICIZATION	0.053	0.013	0.067	0.193	0.337

Figs. 8.4a, b represent the figures graphically:

FIGURE 8.4: Unaccusative verbs, mean auxiliary preferences (ESSERE - AVERE) (BEG = beginner; INT = intermediate; ADV = advanced; ENNS = English near-native, INS = Italian native)

(a)



(b)

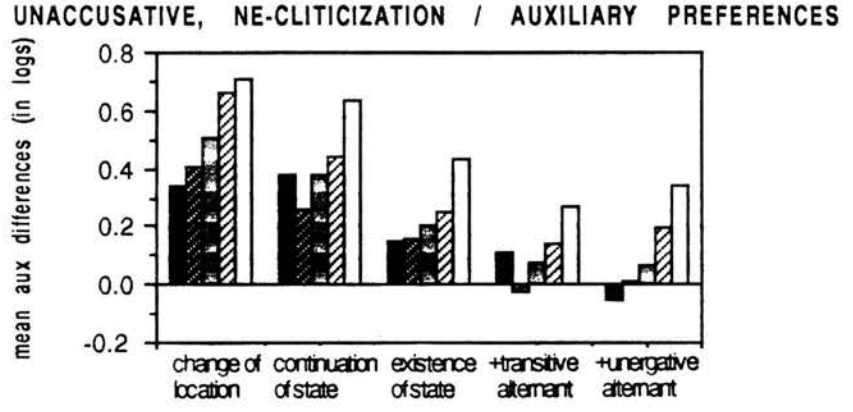


Fig. 8.4a, representing the mean differences in the judgments obtained for basic sentences, illustrates two tendencies in the data.

First, all subject groups - irrespective of proficiency level - distinguish among different categories of unaccusative verbs in a way consistent with the Unaccusative Hierarchy: this means that greater differences between correct ESSERE and incorrect AVERE are found for change-of-location verbs and smallest differences are found for +unergative alternant verbs. A highly significant main effect of semantic category is given by the ANOVA (F_1 (4, 620) = 73.13, $p < .0001$, F_2 (4, 30) = 6.37, $p < .0001$, MinF' (4, 35) = 5.86, $p < .01$).

Second, native speakers have overall higher auxiliary differences than any of the non-native groups. Non-native differences gradually approximate native differences. The main effect for proficiency level has F_1 (4,, 155) = 28.91, $p < .0001$, F_2 (4, 120) = 55.03, $p < .0001$, Min F' (4, 266) = 18.95, and proficiency level does not interact with semantic category. There is, however, a significant interaction of proficiency level, semantic category and word order, F_1 (16, 620) = 2.82, $p < .006$.

These results indicate that both sensitivity to auxiliary selection and discrimination among types of unaccusative verbs develop with proficiency in Italian. More evidence in favour of this interpretation comes from the pattern of differences obtained by applying the Tukey tests to the comparisons among verb categories. There is an increase in the overall number of significant differences at higher proficiency levels. The differences found at lower proficiency level indicate discrimination between core and peripheral unaccusative categories. More sophisticated knowledge of the language entails increased sensitivity to the other distinctions represented along the hierarchy.

At the beginners' level, there are differences concerning both the basic and the Ne-cliticization word orders. For basic sentences (see Fig. 8.4a), auxiliary differences are significantly higher for change-of-location verbs than for [+unergative alternant] verbs (q = 10.58). For Ne-cliticization sentences (see Fig. 8.4b), there are differences between change-of-location and [+unergative alternant] verbs (q = 8.94) and between continuation-of-state and [+unergative alternant] verbs (q = 13.46): in both cases, verb categories closer to the core of the unaccusative hierarchy are associated with larger auxiliary differences.
At the intermediate level, differences are found between change-of-location verbs and [+unergative alternant] verbs in both basic sentences (q = 11.33) and in Ne-cliticization sentences (q = 12.15), as well as between continuation-of-state and [+transitive alternant] verbs in Ne-cliticization word order (q = 8.92).

The differences multiply at the higher proficiency levels, with a single pattern for advanced learners, near-native speakers and native speakers.
At the advanced level, for basic sentences the auxiliary difference of change-of-location verbs is significantly larger than the auxiliary difference of [+unergative alternant] verbs (q = 8.96), [+transitive alternant] verbs (q = 9.47) and existence-of-state verbs (q = 9.16). For Ne-cliticization sentences, there are again significant differences, in the same direction, between change-of-location verbs and [+unergative alternant] ones (q = 13.62), [+transitive alternant] verbs (q = 13.45), and existence-of-state verbs (q = 9.31); also, between

continuation-of-state and [+unergative alternant] verbs (q = 9.61) and between continuation-of-state and [+transitive alternant] verbs (q = 9.44).

At the near-native level, change-of-location verbs are associated, in both word orders, with larger auxiliary differences than those obtained for [+unergative alternant] verbs (basic: q = 13.53; Ne-cliticization: q = 14.62), [+transitive alternant] verbs (basic: q = 12.05; Ne-cliticization: q = 16.26), and existence-of-state verbs (basic: q = 9.56; Ne-cliticization: q = 13). For Ne-cliticization sentences only, continuation-of-state verbs obtain a larger auxiliary difference than [+transitive alternant] verbs (q = 9.41).

At the native level, one again finds stronger preferences, in both word orders, for change-of-location verbs than for [+unergative alternant] verbs (basic: q =12.43; Ne-cliticization: q = 11.4), [+transitive alternant] verbs (basic: q = 10.67; Ne-cliticization: q = 13.59), and existence-of-state verbs (basic: q = 9.18; Ne-cliticization: q = 9.47). There are also stronger preferences for continuation-of-state verbs than both [+unergative alternant] and [+transitive alternant] verbs, but only in Ne-cliticization word order (q = 9.22 and q = 11.41, respectively).

Fig.8.4b presents a similar pattern for Ne-cliticization sentences. Here, the differences between auxiliaries are on the whole higher than those for basic sentences, particularly at lower proficiency levels. This is reflected by a small but significant main effect of word order, F_1 (1, 155) = 4.66), $p < .03$, and by an interaction of word order and proficiency level, F_1 (4, 155) = 3.77, $p < .006$, F_2 (4, 120) = 2.44, $p < .05$.

Nevertheless, the size of differences still tends to be consistent with the position of verbs along the hierarchy. The most important difference between basic and Ne-cliticization sentences can be noticed with respect to the relative order of the two most peripheral verb categories: [+unergative alternant] verbs are generally judged slightly more acceptable with ESSERE than [+transitive alternant] verbs. This explains a significant interaction in the ANOVA of semantic category and word order, $F_1(4, 620) = 3.79$, $p < .005$. These differences, however, are not statistically significant: the Tukey tests applied to the BASIC vs NE-CL comparisons give no differences, which clearly suggests that, overall, auxiliary selection and Ne-cliticization are judged similarly by all subjects.

The overall pattern of inter-group differences is consistent with the tendencies outlined so far. The main boundary is again between learners and native/near-native speakers. The greatest number of significant differences is found for change-of-location verbs, both in basic and Ne-cliticization sentences. Native speakers have significantly higher auxiliary differences than beginners (basic: q = 16.03; Ne-cliticization: q = 11.37), intermediate learners (basic: q = 12.63; Ne-cliticization: q = 9.68), and advanced learners, but only in basic

sentences (q = 11.75). Near-natives differ in the same way from beginners (basic: q = 11.63; Ne-cliticization: q = 10.07).

For continuation-of-state verbs, there are stronger preferences are expressed by natives and weaker preferences by beginners (basic: q = 15.25), intermediates (basic: q = 13.08; Ne-cliticization: q = 11.66), advanced learners (basic: q = 9.26). Near-native speakers' discrimination between auxiliaries is significantly better than beginners', but only with respect to basic sentences (q = 10.47).

For existence-of-state unaccusatives, native speakers have stronger preferences than beginners with respect to basic sentences (q = 13.06), and intermediate learners have weaker preferences than both native and near-native speakers with respect to basic sentences (q = 10.19 and q = 15.97, respectively).

Unaccusative verbs belonging to the [+transitive alternant] category were judged differently, in basic sentences, by native speakers and beginners (q = 13.71), native speakers and intermediate learners (q = 10.35), and native speakers and advanced learners (q = 10.55). With respect to Ne-cliticization sentences, the only significant difference is between intermediate learners and native speakers (q = 9.17).

8.2.3 Summary of findings

The acceptability judgments given by informants on unaccusative verbs support the prediction that the degree of determinacy of auxiliary selection in basic sentences depends on the semantic type of unaccusative verb. As in the case of unergatives, the class of unaccusatives is not homogeneous in the linguistic intuitions of native Italians. Change-of-location verbs are the core, or prototypical type, as shown by the fact that they are the least acceptable with the auxiliary AVERE; [+unergative alternant] verbs are the most peripheral type of unaccusatives, which are the most acceptable with AVERE. The other semantic types of unaccusatives - continuation-of-state, existence-of-state, and [+transitive alternant] - lie between the two extremes of the hierarchy in the predicted order (even though the differences between intermediate verb types are not always statistically significant).

The development of non-native intuitions provides further evidence for the reality of the unaccusative hierarchy. Knowledge of auxiliary selection is first manifested with change-of-location verbs, and then extends to the other verb categories in the predicted order.

Contrary to the hypothesis that Ne-cliticization would be immune to semantic distinctions both in synchronic and in developmental terms, the judgments on Ne-cliticization sentences present a similar pattern to the judgments on basic sentences. As in the case of unergative verbs, the unaccusative results disconfirm the hypothesis that the syntactic reflexes of unaccusativity/unergativity distinction are impenetrable to semantic characterizations. The

development of non-native linguistic intuitions on Ne-cliticization and on basic sentences follow similar paths , thus providing further support to the idea that syntactic phenomena with semantic correlates do not pose particular learnability problems.

8.3 Restructuring constructions

The final section of this chapter is concerned with the acceptability judgements obtained on auxiliary selection in the set of constructions that we have grouped under the general term of 'restructuring', even though this set consists of both restructured and non-restructured sentences. Unlike auxiliary choice in the present perfect tense, the behaviour of auxiliaries in these constructions appears to be a function of well-defined syntactic properties induced by a restricted number of verbs, and is not characterizable in terms of semantic generalizations. Depending on the syntactic configuration that the verb enters into, both auxiliaries are allowed (as in basic restructured sentences, both with and without pronominal arguments), or only ESSERE is allowed (as under clitic-climbing), or only AVERE is required (as under clefting). On a purely syntactic level, restructuring constructions are marked or peripheral, compared to more central cases. The knowledge required to select the appropriate auxiliary is therefore of a complex nature, and the evidence necessary to acquire such knowledge is variable and, in some cases, scanty.

The hypotheses for restructuring constructions predicted that:

(a) the judgments of native speakers about auxiliary selection in restructuring sentences would be less determinate than their intuitions about auxiliary selection with the present perfect tense, but nevertheless uniform: grammatical auxiliaries would be preferred to ungrammatical auxiliaries, irrespective of the particular construction or of the inducing verb. Bearing in mind that ESSERE is an indication of restructuring, whereas AVERE indicates non-restructuring, it was predicted that in native speakers' judgments there would be: (i) no difference between raising and control verbs; (ii) no difference between ESSERE and AVERE versions of basic sentences; (iii) no difference between ESSERE and AVERE versions of sentences containing non-moved pronominal arguments; (iv) a clear difference between correct ESSERE and incorrect AVERE versions of sentences containing clitic-climbing; (v) a clear difference between correct AVERE and incorrect ESSERE versions of sentences exhibiting clefting.

(b) the judgments of non-native speakers would NOT show a pattern of gradual approximation to native judgments. Non-natives would find it problematic to acquire the

syntactic principles responsible for auxiliary choice in restructuring constructions; they would therefore tend to be indeterminate up to a certain level, and possibly even at very high proficiency levels.

These prediction were analysed statistically by means of the following tests:

(a) a pair of four-way repeated-measures ANOVAs (proficiency level x verb type (raising vs control) x syntactic category (restructured vs non-restructured) x word order);

(b) Min F' values calculated on pairs of significant F_1 and F_2 values;

(c) Post-hoc Tukey tests of significance for differences between means applied on the F_1 means . The critical value of q for all comparisons is 6.54.

8.3.1 All variables

Table 8.5 shows the geometric means of the acceptability scores obtained for restructuring verbs (notice that the distinction between AVERE and ESSERE corresponds to the distinction between 'non-restructured' and 'restructured', and will be used in place of the latter):

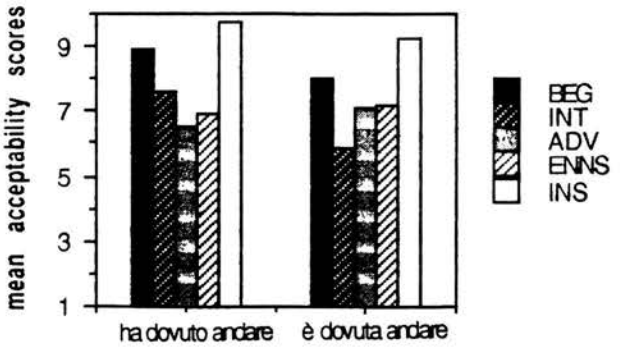
TABLE 8.5: Mean acceptability judgments on Restructuring verbs (BEG = beginner; INT = intermediate; ADV = advanced; ENNS = English near-native, INS = Italian native)

	BEG	INT	ADV	ENNS	INS
RAISING, BASIC, NON-RESTRUCTURED	8.919	7.587	6.547	6.977	9.749
RAISING, CLEFTING, NON-RESTRUCTURED	4.088	3.471	4.088	3.644	5.817
RAISING, CLIT-CLIMB, NON-RESTRUCTURED	6.537	6.430	6.217	6.623	3.143
RAISING, NO-CLIT-MOV, NON-RESTRUCTURED	6.337	5.378	6.352	6.211	8.779
RAISING, BASIC, RESTRUCTURED	7.980	5.859	7.069	7.232	9.260
RAISING, CLEFTING, RESTRUCTURED	3.806	3.586	3.510	3.075	3.039
RAISING, CLIT-CLIMB, RESTRUCTURED	6.107	5.239	6.024	6.286	8.587
RAISING, NO-CLIT-MOV., RESTRUCTURED	6.408	4.977	6.086	6.784	8.159
CONTROL, BASIC, NON-RESTRUCTURED	8.902	7.890	8.151	8.816	9.265
CONTROL, CLEFTING, NON-RESTRUCTURED	3.960	3.612	4.242	3.090	5.466
CONTROL, CLIT-CLIMB, NON-RESTRUCTURED	7.015	6.662	7.123	7.704	3.443
CONTROL, NO-CLIT-MOV, NON-RESTRUCTURED	7.198	6.447	6.799	7.610	8.279
CONTROL, BASIC, RESTRUCTURED	6.245	5.361	6.506	5.876	7.081
CONTROL, CLEFTING, RESTRUCTURED	4.186	3.061	3.516	2.974	3.062
CONTROL, CLIT-CLIMB, RESTRUCTURED	6.766	4.713	5.656	6.259	6.768
CONTROL, NO-CLIT-MOV, RESTRUCTURED	6.742	5.306	5.741	6.028	6.294

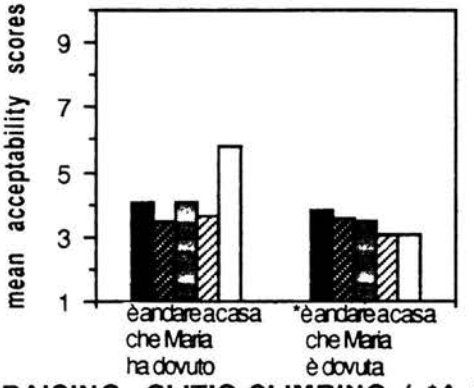
These results are represented graphically in Figs. 8.5a-d for Raising verbs and in Figs.8.6a-d for Control verbs:

FIGURE 8.5: Restructuring constructions with Raising verbs (BEG = beginner; INT = intermediate; ADV = advanced; ENNS = English near-native, INS = Italian native)

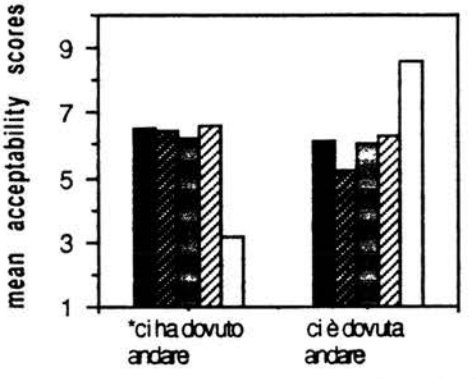
(a) RAISING, BASIC / A VS E



(b) RAISING, CLEFTING / A VS *E



(c) RAISING, CLITIC-CLIMBING / *A VS E



(d) RAISING, NO-CLITIC-MOVEMENT / A VS E

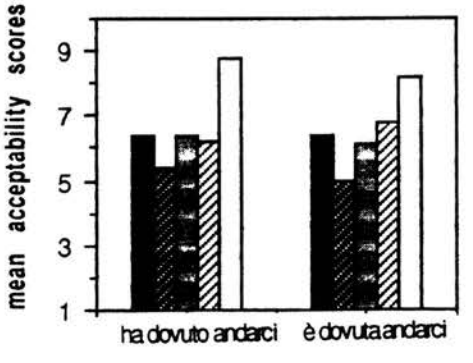
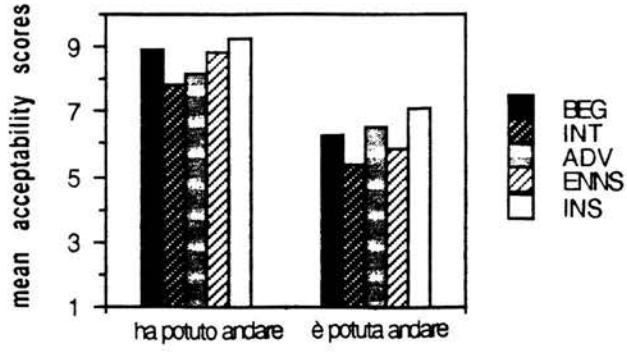
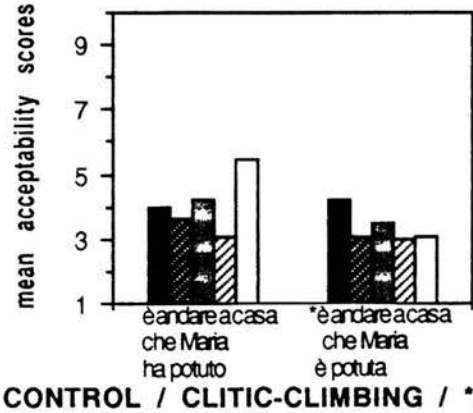


FIGURE 8.6: Restructuring constructions with Control verbs (BEG = beginner; INT = intermediate; ADV = advanced; ENNS = English near-native, INS = Italian native)

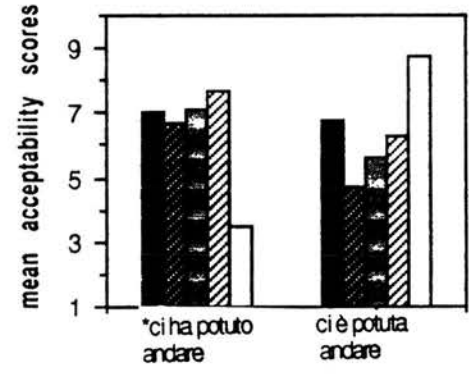
(a) CONTROL / BASIC / A VS E



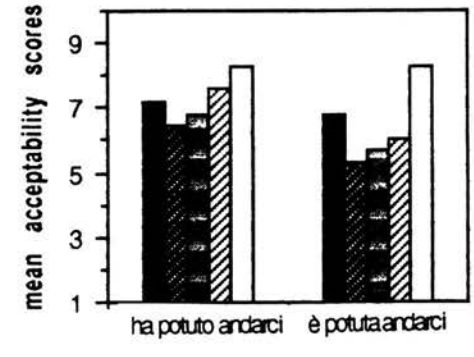
(b) CONTROL / CLEFTING / A VS *E



(c) CONTROL / CLITIC-CLIMBING / *A VS E



(d) CONTROL / NO-CLITIC-MOVEMENT / A VS E



Let us first examine the pattern of acceptability judgments on Raising verbs, shown in Fig.8.5. The most conspicuous fact exhibited by these graphs is a clear difference between native and non-native speakers: while native speakers give determinate judgments that distinguish between acceptable and unacceptable auxiliary choices, non-native speakers give similar, non-discriminating judgments on most sentences. Unlike the judgments on unergative and unaccusative verbs, judgments on restructuring verbs do not exhibit any progression across proficiency levels. Statistically, in fact, the main effect of proficiency level does not reach significance.

There is, however, an interaction of proficiency level and word order, F_1 (12, 465) = 3.92, $p < .0001$, F_2 (12, 96) = 3.00, $p < .0013$. It is only natives who clearly accept both auxiliaries (and therefore recognize the grammaticality of optional restructuring) in basic sentences and sentences without clitic movement, as can be seen in Fig. 5a; non-natives judge both auxiliaries in a similar way but their judgments do not get beyond the indeterminate middle range. Even more conspicuously, it is only natives who distinguish between correct and incorrect auxiliaries in sentences including clefting and clitic-climbing, and therefore discriminate between possible and impossible cases for restructuring. Non-natives do not make any discrimination between auxiliaries for these sentence types.
The ANOVA gives significant main effects not only for both syntactic category, F_1 (1, 155) = 13.96, $p < .0003$, F_2(1, 24) = 10.99, $p < .0029$, Min F' (1, 70) = 6.15, $p < .05$, and word order, F_1 (3, 465) = 150.81, $p < .0001$, F_2 (3, 24) = 73.26, $p < .0001$, Min F' (3, 52) = 49.31, $p < .01$, but also for the interaction of syntactic category, word order and proficiency level, F_1(12, 465) = 12.27, $p < .0001$, F_2(12, 96) = 6.17, $p < .0001$, Min F' (12, 206) = 4.11, $p < .01$.

The Tukey tests conducted on the comparisons among proficiency levels also show that all non-native subjects do not distinguish between ESSERE and AVERE in sentences exhibiting clitic-climbing. While there are no significant differences among non-native groups, the most conspicuous differences between native and non-native judgments concerns clitic-climbing, which is consistently accepted by non-native speakers with both auxiliaries. In non-restructured sentences, native speakers find clitic-climbing significantly less acceptable than beginners (raising: q = 10.26; control: q = 9.68), intermediate learners (raising: q = 9.74; control: q = 9.98), advanced learners (raising: q = 9.28; control: q = 9.89), and near-natives (raising: q = 10.14; control: q = 10.95).
In restructured sentences, native judgments are significantly higher than intermediate learners' judgments on clitic-climbing (raising: q = 6.72; control: q = 8.44), and no-clitic-movement (raising: q = 6.72).

The disparity between native and non-native speakers is further confirmed by the pattern of differences obtained by applying the Tukey test to the comparisons between auxiliaries (recall that these correspond to the presence or absence of syntactic restructuring). While no significant differences are found at any non-native levels, at the native level one finds the expected significant differences, for both raising and control verbs, between AVERE and ESSERE in clefted sentences, (raising: q = 8.83; control: q = 7.88), and in sentences with clitic-climbing (raising: q = 13.67; control: q =12.71). Recall that these are precisely the cases in which only one auxiliary is grammatical: AVERE with clefting, and ESSERE with clitic-climbing. There are no differences between auxiliaries for sentences in basic word order, both with full and with pronominal arguments: this, again, confirms the predictions, for they are both acceptable.

The overall picture provided by the differences among word orders confirms the major gap between the judgments of non-natives and those of natives, but is complicated by the particular difficulty of clefted sentences compared to the other constructions that informants were asked to judge. In non-native judgments, the only significant word order difference separates clefted sentences from the others: clefting is consistently given lower acceptability scores than any of the other word orders, irrespective of the auxiliary with which it appears (this can be seen in both Fig. 8.5b and Fig. 8.6b). This effect is clearly due to the oddity and 'unnaturalness' of clefted sentences, and it therefore has little theoretical relevance. Thus, in the judgments of beginners on non-restructured sentences, one finds significant differences between clefted sentences and sentences in basic word order (raising: q = 10.6; control: q = 11.02) and clitic-climbing (raising: q = 6.68; control: q = 7.78), whereas in judgments on restructured sentences clefting differs from basic (raising: q = 10.06), clitic-climbing (raising: q = 6.61; control: q = 7.29), and no-clitic-movement (raising: q = 7.08; control: q = 6.48).

In the judgments of intermediate learners, clefting significantly differs from basic in both non-restructured (raising: q = 8.64; control: q= 10.63) and in restructured sentences (raising: q = 6.61; control: q = 7.62).

In the judgments of advanced learners on non-restructured sentences, clefting is significantly different from basic (raising: q = 9.52), clitic-climbing (raising: q = 7.35; control: q = 7.05), and no-clitic-movement (raising: q = 7.49; control: q = 6.62); in restructured sentences, clefting differs from basic (raising: q = 9.52; control: q = 8.37), clitic-climbing (raising: q = 7.35), and no-clitic-movement (raising: q = 7.49).

Differences become more marked at the near-native level, where , in non-restructured sentences, clefting is judged differently from basic (raising: q = 8.83; control: q = 10.76), clitic-climbing (raising: q = 8.13; control: q = 12.42), and no-clitic-movement (raising: q = 7.25;

control: q = 12.26). In restructured sentences, clefting differs from basic (raising: q = 11.63; control: q = 9.26), clitic-climbing (raising: q = 9.73; control: q = 10.12), and no-clitic-movement (raising: q = 10.76; control: q = 9.63).

One should not conclude, on the basis of these results, that non-native informants have knowledge of auxiliary choice with clefting because, as we have just seen, they do not make any significant difference between auxiliaries for any sentence type.

Finally, the judgments of native speakers exhibit the complete pattern of expected differences. In non-restructured sentences, clitic-climbing is judged differently from basic word order (raising: q = 15.39; control: q = 13.46), from clefting (raising: q = 8.37; control: q = 6.51), and from no-clitic-movement (raising: q = 13,97; control: q = 11.93). In restructured sentences, judgments on clefting are different from those on basic word order (raising: q = 15.15; control: q = 11.4), clitic-climbing (raising: q = 14.13; control: q = 14.3), and no-clitic-movement (raising: q = 13.43; control: q = 13.55). However, judgments on clefting in non-restructured sentences are significantly lower than judgments on basic word order (raising: q = 7.02; control: q = 7.17), which suggest that the judgments of native speakers are sensitive to both the choice of auxiliary and the markedness of clefted sentences.

The pattern of raising verbs is, to a large extent, duplicated by the graphs in Fig.8.6a-d, representing judgments on control verbs. This supports the prediction that Raising and Control verbs would be equivalent with respect to the syntactic properties subsumed under restructuring.

The only difference between raising and control verbs seems to lie in the judgments given by non-native speakers on sentences with basic word order and sentences with no-clitic-movement, which appear to be slightly higher for control verbs in the non-restructured condition than for the corresponding raising verbs. This justifies the significant interaction obtained in the ANOVA of verb type and syntactic category, F_1 (1, 155) = 11.06, $p < .001$, $F_2(1, 24) = 4.41$, $p < .046$. However, the Tukey tests reveal no significant differences between judgments on raising and judgments on control verbs for any of the subject groups.

8.3.2 Auxiliary preferences

The informants' knowledge of auxiliary selection under restructuring was further tested by analysing the differences between auxiliaries for each of the constructions investigated. Recall that auxiliary ESSERE indicates restructuring, whereas auxiliary AVERE indicates absence of restructuring. Given the individual behaviour of the four constructions vis-à-vis

restructuring (i.e. the fact that basic and no-clitic-movement constructions admit both auxiliaries, clefting allows only AVERE , and clitic-climbing allows only ESSERE , the differences were calculated as follows:

- BASIC: AVERE - ESSERE (expected difference = 0)
- CLEFTING: AVERE - ESSERE (expected difference > 0)
- CLITIC-CLIMBING: ESSERE - AVERE (expected difference> 0)
- NO CLITIC MOVEMENT: AVERE - ESSERE (expected difference = 0)

The prediction was that substantial differences would be obtained in native judgments for clefting and clitic-climbing, whereas differences close to zero would be obtained for basic and no-clitic-movement constructions. For non-native judgments, it was predicted that all differences would be close to zero.

These predictions were tested via the following statistical procedures:

(a) a pair of three-way repeated measures ANOVAs (proficiency level x verb type x word order);

(b) Min F' values for all pairs of significant F_1 and F_2 values;

(c) Post-hoc Tukey tests of significance were applied to the relevant comparisons between F_1 means. The critical value of q at $p < .01$ for all comparisons is 9.13.

The mean differences obtained are reported (in logarithmic form) in Table 8.6. Again, positive numbers mean preference for AVERE, except for clitic-climbing, where they mean preference for ESSERE; negative numbers mean a preference for ESSERE (or for AVERE, in the case of clitic-climbing). Larger numbers express stronger preferences.

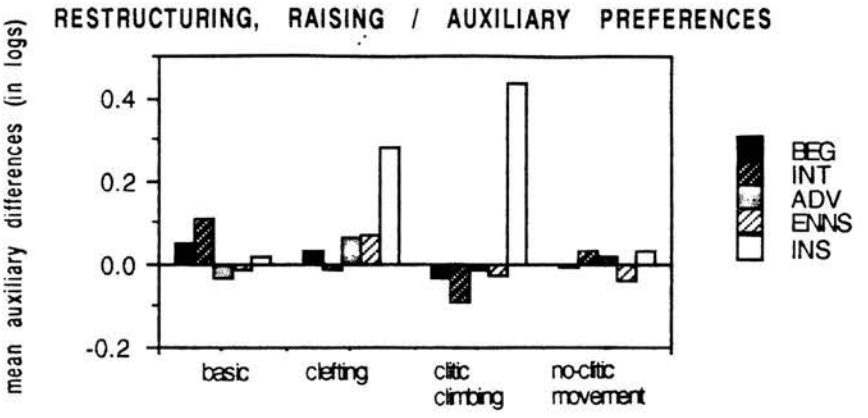
TABLE 8.6: Restructuring verbs, mean auxiliary differences (correct auxiliary - incorrect auxiliary) (BEG = beginner; INT = intermediate; ADV = advanced; ENNS = English near-native, INS = Italian native)

	BEG	INT	ADV	ENNS	INS
RAISING, BASIC	0.048	0.112	-0.033	-0.016	0.022
RAISING, CLEFTING	0.031	-0.014	0.066	0.074	0.282
RAISING, CLITIC-CLIMBING	-0.030	-0.089	-0.014	-0.023	0.436
RAISING, NO-CLITIC-MOVEMENT	-0.005	0.034	0.019	-0.038	0.032
CONTROL, BASIC	0.154	0.168	0.003	0.064	0.117
CONTROL, CLEFTING	-0.024	0.072	0.082	0.017	0.252
CONTROL, CLITIC-CLIMBING	-0.016	-0.150	-0.100	-0.090	0.406
CONTROL, NO-CLITIC-MOVEMENT	0.028	0.085	0.073	0.101	0.001

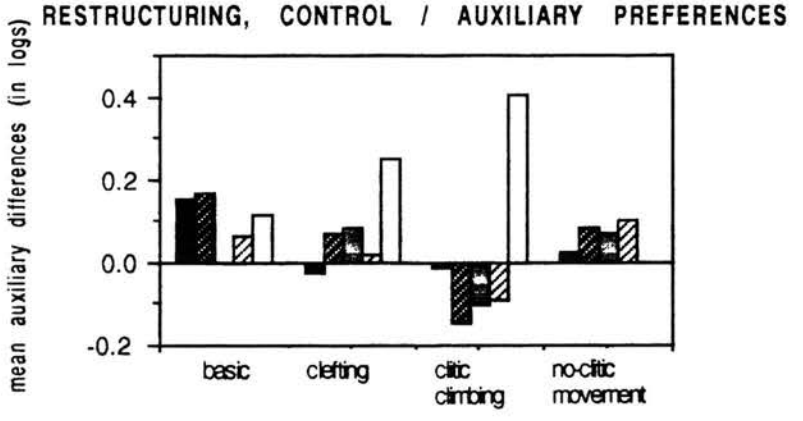
The graphic representation of these figures is given in Fig.8.7a,b.

FIGURE 8.7: Restructuring verbs, mean auxiliary differences (correct auxiliary - incorrect auxiliary) for (a) Raising and (b) Control (BEG = beginner; INT = intermediate; ADV = advanced; ENNS = English near-native, INS = Italian native)

(a)



(b)



The two graphs representing raising (Fig. 8.7a) and control verbs (Fig. 8.7b) look similar, except for marginally larger differences obtained for non-native judgments on control verbs: these generally indicate more marked preferences for the AVERE auxiliary. The ANOVA, however, gives no significant main effect of verb type.

In Fig. 8.7a, b native speakers clearly differ from non-native speakers in having much stronger auxiliary preferences for clitic-climbing and, to a lesser extent, for clefting. There is a main effect of proficiency level, F_1 (4, 155) = 11.52, $p < .0001$, F_2 (4, 24) = 8.72, $p < .0001$, Min F' (4, 68) = 4.96, $p < .01$, and a significant interaction of word order and proficiency level, F_1 (12, 465) = 5.86, $p < .0001$, F_2 (12, 96) = 3.71, $p < .0001$, Min F' (12, 235) = 2.28, $p < .01$. Tukey tests show no significant differences among non-native subject groups. Differences between natives and non-natives concern clitic-climbing, where natives' auxiliary differences are significantly higher than those of beginners (raising: q = 14.6; control: q = 13.2), intermediate learners (raising: q = 16.46; control: q = 17.42), advanced learners (raising: q = 14.25; control: q = 15.85), and near-natives (raising: q = 14.37; control: q = 15.54). There are also significant differences between the stronger native preferences in auxiliary acceptability in clefted sentences with raising verbs and the corresponding weaker preferences of intermediate learners (q = 9.27).
The general pattern suggests that increasing knowledge of the foreign language involves growing awareness of unacceptability, rather than of acceptability, as we saw in Chapter 7.

Finally, there is a noticeable uniformity in non-native auxiliary differences for different sentence types. Tukey test show no significant differences between word orders for judgments given by non-native informants. At the native level, however, the mean auxiliary preference for judgments obtained on clitic-climbing is significantly greater than that obtained for basic sentences (raising: q = 12.97; control: q = 9.15), and for no-clitic-movement (raising: q = 12.68; control: q = 12.71).

8.3.3 A comparison with Card-Sorting

It is interesting at this point to compare the results of Magnitude Estimation with the corresponding ones of Card-Sorting. Although substantially similar to the patterns we have just discussed, the judgments obtained via Card-Sorting show (a) a more marked difference between Raising and Control verbs, and (b) a closer resemblance between the judgments of near-native speakers and those of native speakers.
The ANOVAs run on all variables (see Tables B39 and B51 in Appendix B) yield significant interactions of verb type and syntactic category, F_1 (1, 155) = 19.22, $p < .0001$, F_2 (1, 24) =

8.69, p <.007, Min F' (1, 49) = 5.98, p <.05, verb type and word order, F_1 (3, 465) = 4.73, p <.003, verb type, syntactic category and word order, F_1 (3, 465) = 5.75, p <.0007, and verb type, syntactic category, word order and proficiency level, F_1 (12, 465) = 3.08, p <.0003.
The ANOVAs on auxiliary preferences give a significant main effect of verb type, F_1 (1, 155) = 15.36, p <.0001, F_2 (1, 18) = 9.15, p <.007, Min F' (1, 44) = 5.73, p <.05, as well as significant interactions of verb type and proficiency level, F_1 (4, 155) = 3.76, p <.006, F_2 (4, 72) = 2.74, p <.04, and verb type and word order, F_1 (3, 465) = 10.13, p <.0001.
Tukey tests applied to comparisons among groups reveal no significant differences between near-native and native judgments for any syntactic category, word order, or verb type, unlike the tests appplied to the corresponding judgments obtained through Magnitude Estimation .

8.3.4 Summary of findings

The judgments on restructuring constructions confirm the prediction that Raising and Control verbs are equivalent in their syntactic properties. Native speakers uniformly accept the AVERE ---> ESSERE change as a syntactic reflex of restructuring in both the optional cases (basic and no-clitic-movement) and the obligatory cases (clitic-climbing), and they reject it when restructuring is disallowed (clefting).
Further, the data also support the prediction that auxiliary change under restructuring, as a purely syntactic phenomenon of relative marginality, is more difficult to acquire. Consistent with this hypothesis, non-native judgments show no evidence of development in the direction of native acceptability values. Even at the near-native level, there is indeterminacy of judgment.
Interestingly, the indeterminacy of near-native judgments is less noticeable in the Card-Sorting results: in the untimed condition of this method, near-natives produce judgments that are similar to those of native speakers.

CHAPTER 9
RESULTS (3): ULTIMATE ATTAINMENT

9.0 Introduction

This chapter analyses the knowledge of auxiliary selection attained by near-native speakers of Italian. It compares the acceptability judgments expressed by English near-natives and French near-natives with the judgments of native Italian speakers. This comparisons will provide evidence for or against the two experimental hypotheses raised in Chapter 4, which predicted that:

(a) the overall pattern of near-native intuitions would be different from the pattern of native intuitions;

(b) in general, near-native intuitions would be more similar to native intuitions with respect to the semantic properties of auxiliary selection than to the syntactic properties because the former should be easier to acquire for second language learners than the latter;

(c) for syntactic properties, the intuitions of French near-natives would be more similar to those of native speakers than the intuitions of English-near-natives; this is because English presents no instantiation of the properties exhibited by Italian, whereas French presents partial instantiation of such properties.

This chapter will have the same structure as Chapter 7: it will deal first with unergative verbs, then with unaccusative verbs, and finally with restructuring verbs.

9.1 Unergative verbs

For this class of verbs, it was predicted that:

(a) both native and near-native speakers would not respond uniformly in their judgments on unergative verbs: the degree of determinacy of judgments on auxiliary selection in basic sentences would be related to the type of unergative verb, according to the following hierarchy:

non-motional > motional > [+unaccusative alternant]

(b) both native and near-native speakers would tend to reject Ne-cliticized sentences containing unergative verbs, irrespective of semantic type and of auxiliary , because Ne-cliticization should not be sensitive to lexical-semantic characterizations;

(c) English and French near-native speakers would provide similar acceptability judgments in basic sentences, but French near-natives' judgments on Ne-cliticization sentences would be closer to native Italian judgments than those of English near-natives, because of the differences between the two languages.

In order to test these predictions, the data were subjected to the following statistical treatments:

(a) a pair of four-way ANOVAs with repeated measures (native language x semantic category x auxiliary x word order), both by subjects and by materials;
(b) Min F' values for pairs of significant F_1 and F_2 values;
(c) post-hoc Tukey tests between pairs of F_1 means. The critical value of q at $p < .01$ for all comparison is 11.89.

9.1.1 All variables

The geometric means of the acceptability judgments obtained from the three subject groups are reported in Table 9.1.

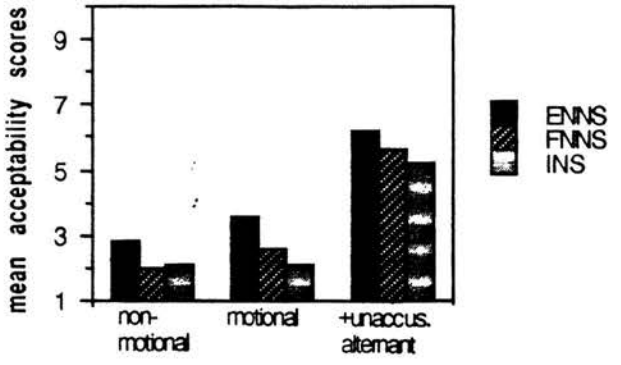
TABLE 9.1: Mean acceptability judgments on unergative verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).

	ENNS	FNNS	INS
NON-MOTIONAL, BASIC, ESSERE	2.810	2.025	2.067
NON-MOTIONAL, BASIC, AVERE	10.716	10.286	9.841
NON-MOTIONAL, NE-CLITICIZATION, ESSERE	2.779	2.313	2.054
NON-MOTIONAL, NE-CLITICIZATION, AVERE	7.943	4.236	3.105
MOTIONAL, BASIC, ESSERE	3.555	2.536	2.087
MOTIONAL, BASIC, AVERE	9.940	10.581	9.811
MOTIONAL, NE-CLITICIZATION, ESSERE	4.320	4.785	2.913
MOTIONAL, NE-CLITICIZATION, AVERE	6.838	5.100	3.216
[+UNACCUSATIVE ALTERNANT], BASIC, ESSERE	6.241	5.596	5.206
[+UNACCUSATIVE ALTERNANT], BASIC, AVERE	6.420	7.871	7.542
[+UNACCUSATIVE ALTERNANT], NE-CLITICIZATION, ESSERE	6.329	4.757	5.515
[+UNACCUSATIVE ALTERNANT], NE-CLITICIZATION, AVERE	5.636	4.101	3.079

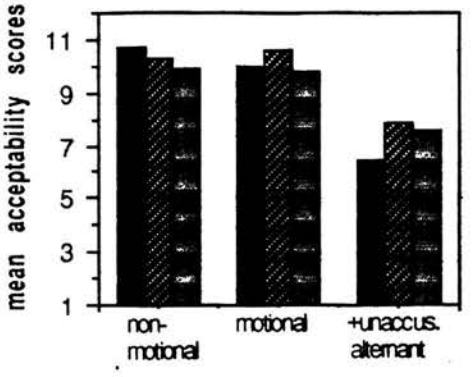
The data are also represented graphically in Figures 9.1a-d.

FIGURE 9.1: Mean acceptability judgments on unergative verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers)

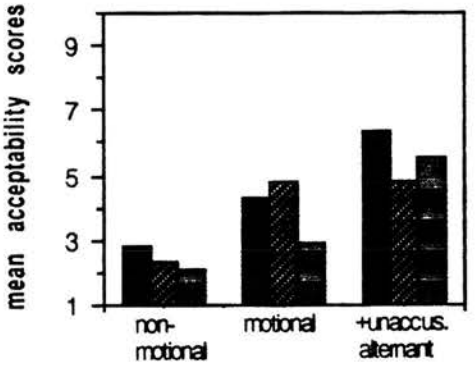
(a) UNERGATIVE, BASIC / *ESSERE



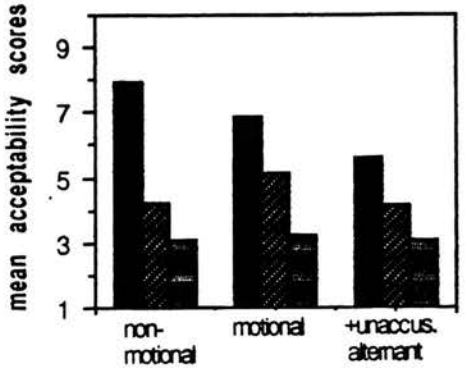
(b) UNERGATIVE, BASIC / AVERE



(c) UNERGATIVE, *NE-CLITICIZATION / *ESSERE



(d) UNERGATIVE, *NE-CLITICIZATION / *AVERE



The first overall impression given by the graphs in Figures 9.1a-d is that native and near-native judgments are not identical. The main effect of native language is in fact significant, F_1 (2, 77) = 4,68, p <.01, F_2 (2, 36) = 18.17, p <.0001, Min F' (2, 107) = 3.72, p <.05.

Second, the types of unergative verbs submitted to the informants' judgments are given different responses. There is a significant main effect of semantic category, F_1 (2, 154) = 28.30, p < .0001, F_2 (2, 18) = 16.70, p < .0001, Min F' (2, 37) = 19.35, p <.01. Variation determined by semantic categories is particularly obvious in Figures 9.1a and 9.1c, which show the mean acceptability judgments on basic and Ne-cliticization sentences containing ESSERE (as in the developmental data examined in Chapter 7, sentences with AVERE elicit more uniform responses, independently of verb category). The interaction of semantic category and auxiliary is highly significant, F_1 (35, 154) = 61.47, p <.0001, F_2 (35, 18) = 28.24, p < .0001, Min F' (2, 37) = 19.35, p <.01, and so is the interaction of semantic category, word order and auxiliary, F_1 (2, 154) = 9.44, p <.0001. These patterns of responses are similar across native languages.The interaction of semantic category and native language is significant, F_1 (35, 154) = 2.88, p <.02, suggesting that the three informant groups respond to sentences in a similar way.

Evidence for the reality of the distinctions among unergative verbs is also provided by Tukey tests of significance on comparisons among verb categories. All subject groups distinguish between non-motional verbs and [+unaccusative alternant] verbs presented with ESSERE, both in basic (English: q = 20.05; French: q = 25.55; Italian: q = 22.86) and in Ne-cliticization sentences (English: q = 20.69; French: q = 18.13; Italians: q = 21.92): informants judge [+unaccusative alternant] verbs significantly more acceptable with ESSERE than non-motional verbs. All groups also differentiate between motional unergatives and [+unaccusative alternant] ones, but only in basic sentences (English: q = 14.14; French : q = 19.9; Italians : q = 19.56): again, [+unaccusative alternant] verbs are perceived as more acceptable with ESSERE than motional verbs.
Two differences, however, are peculiar to near-native subjects: as can be seen in Fig. 9.1b , English informants prefer non-motional to [+unaccusative alternant] verbs in basic sentences with AVERE (q = 12.88). French informants (see Fig. 9.1c) prefer motional to non-motional verbs in Ne-cliticization sentences with ESSERE. Native Italians do not make any comparable discrimination.

Third, the pattern of judgments for basic sentences and Ne-cliticization sentences are similar, but only with respect to the ESSERE auxiliary, as can be seen in Figures 9.1a, c. For sentences with AVERE, responses provide two distinct pictures. Predictably, basic

sentences are generally judged acceptable by both native and near-native speakers, although [+unaccusative alternant] verbs are perceived as less acceptable than the other two types. Ne-cliticization sentences, on the other hand, are judged in the same way regardless of verb category by native Italians and French near-natives, although French acceptability ratings are lower than Italian ones (as can be seen in Fig. 9.1d). Verb category, however, does make a difference to English near-natives, who tend to find non-motional verbs the most acceptable and [+unaccusative alternant] the least acceptable. In fact, the only significant subject group difference is between English near-natives' judgments on Ne-cliticized non-motional verbs with AVERE and the French near-natives' and Italian natives' lower judgments of the same sentences ($q = 15.91$ and $q = 14.68$ respectively)

Statistically, the significant main effects of word order, F_1 (1, 77) = 46.12, $p < .0001$, F_2 (1, 18) = 38.46, $p < .0001$, Min F' (1, 52) = 20.97, $p < .01$, and of auxiliary, F_1 (1, 154) = 157.05, $p < .0001$, F_2 (1, 18) = 98.00, $p < .0001$, Min F' (1, 45) = 60.34, $p < .01$, are qualified by a significant interaction of word order and auxiliary, F_1 (1, 77) =91.60, $p < .0001$, F_2(1, 18) = 43.77, $p < .0001$, Min F' (1, 37) = 29.62, $p < .01$, consistent with the influence of auxiliary on the parallel between word orders. The differences between English informants, on the one hand, and French and Italian informants on the other with respect to Ne-cliticization sentences with AVERE is reflected by a significant interaction of word order and native language, F_1(2, 77) = 7.45, $p < .001$, F_2 (2, 36) = 5.13, $p < .01$, and by a three-way interaction of word order, auxiliary and native language, F_1(2, 77) = 9.45, $p < .0002$, F_2(2, 36) = 11.56, $p < .0001$, Min F' (2, 104) = 5.2, $p < .01$.

Tukey tests also show that the judgments of the three informant groups are not equivalent with respect to word order. It was predicted that native speakers would (a) accept basic sentences and reject Ne-cliticization sentences with AVERE, and (b) reject both basic and Ne-cliticization sentences with ESSERE (although semantic categories would determine the strength of the rejection in basic but not in Ne-cliticization sentences). This is precisely what is found: there are no significant differences in native judgments between basic and Ne-cliticization sentences with ESSERE, but acceptability ratings on basic word order are significantly higher than those on Ne-cliticization in sentences containing AVERE, with all the three verb categories (non-motional: $q = 20.88$; motional: $q = 20.03$; [+unaccusative alternant]: $q = 15.72$).

It was also predicted that the judgments of French near-native speakers on Ne-cliticization would be more similar to the judgments of Italian native speakers than those of English near-natives. This is also the case. There are no significant differences between word orders in the judgments of English near-natives, suggesting that Ne-cliticization with unergative verbs

is possible in their knowledge representation : this can be clearly seen in Figures 9.1a, c and 9.1b, d, which show very similar patterns for the two word orders with both auxiliaries. On the other hand, French near-natives - like native Italians - give higher acceptability scores to basic sentences with AVERE than to Ne-cliticized sentences, within all verb categories (non-motional: q = 22.31; motional: q = 18.35; [+unaccusative alternant]: q = 16.39). Unlike native speakers, they also prefer Ne-cliticization sentences containing motional verbs with ESSERE to the equivalent basic sentences (q = 16.39).

The similarity between the judgments of French near-native speakers and those of Italian native speakers is further confirmed by the Tukey tests on the differences between ESSERE and AVERE. All groups significantly prefer AVERE to ESSERE in basic sentences with non-motional verbs (English: q = 33.65; French: q = 40.86; Italian: q = 37.96) and with motional verbs (English: q = 25.85; French: q = 35.92; Italian: q = 34.24). All groups also prefer AVERE in Ne-cliticization sentences with non-motional unergatives (English: q = 26.5; French: q = 15.21; Italian: q = 16.4). No group has any significant difference between auxiliaries with [+unaccusative alternant] verbs, regardless of word order. English near-natives, however, also have a significant preference for AVERE with motional verbs in Ne-cliticization sentences, thus again showing a parallelism between basic and Ne-cliticization word orders in their knowledge of auxiliary selection with unergative verbs.

9.1.2 Auxiliary preferences

To have a more direct indication of the informants' knowledge of auxiliary selection with unergative verbs, statistical analyses were carried out on the mean differences obtained by subtracting judgments on sentences containing the auxiliary ESSERE from judgments on corresponding sentences containing the auxiliary AVERE. These mean auxiliary differences (in logarithmic form) are shown in Table 9.2, where positive numbers indicate a preference for AVERE, negative numbers a preference for ESSERE, and larger differences indicate stronger preferences.

The predictions for these data can be summarized as follows:

(a) all groups would have the strongest (positive) auxiliary preferences with non-motional unergative verbs, and the weakest (not necessarily positive) auxiliary preferences with [+unaccusative alternant] verbs in basic form;

(b) because no Ne-cliticized sentences with unergative verbs should be acceptable in Italian, native speakers would have auxiliary preferences close to zero with Ne-

cliticization sentences, irrespective of verb category. French near-natives would resemble native speakers more closely than English near-natives do.

These hypotheses were subjected to the following statistical analyses:

(a) a pair of three-way repeated-measures ANOVAs (proficiency level x semantic category x word order);

(b) Min F' values were calculated on pairs of significant F_1 and F_2 values;

(c) Tukey tests of significance between F_1 means were applied on the basis of the ANOVA results. The critical value of q at $p < .01$ is 16.8.

The means for the auxiliary preferences expressed by the three subject groups are shown in Table 9.2 (in logarithmic form).

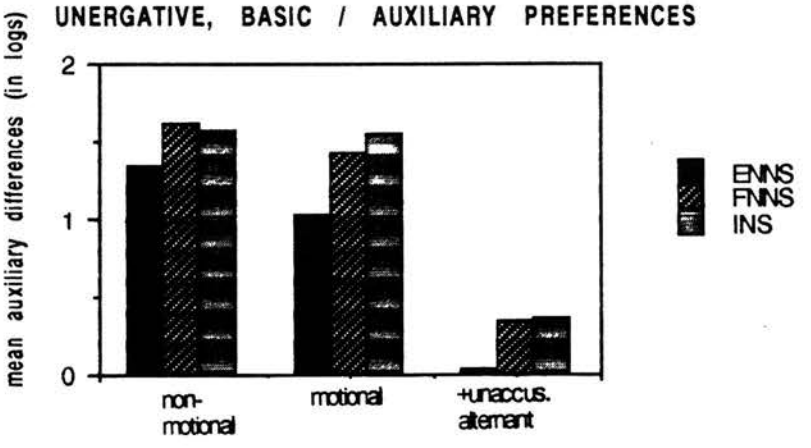
TABLE 9.2: Unergative verbs, mean auxiliary preferences (ENNS = English near-native speakers; FNNS = French near-native speakers; INS = native Italian speakers)

	ENNS	FNNS	INS
NON-MOTIONAL, BASIC	1.339	1.625	1.561
NON-MOTIONAL, NE-CLITICIZATION	1.050	0.605	0.413
MOTIONAL, BASIC	1.028	1.429	1.548
MOTIONAL, NE-CLITICIZATION	0.459	0.064	0.099
[+UNACCUSATIVE ALTERNANT], BASIC	0.028	0.341	0.371
[+UNACCUSATIVE ALTERNANT], NE-CLITICIZATION	-0.116	-0.148	-0.583

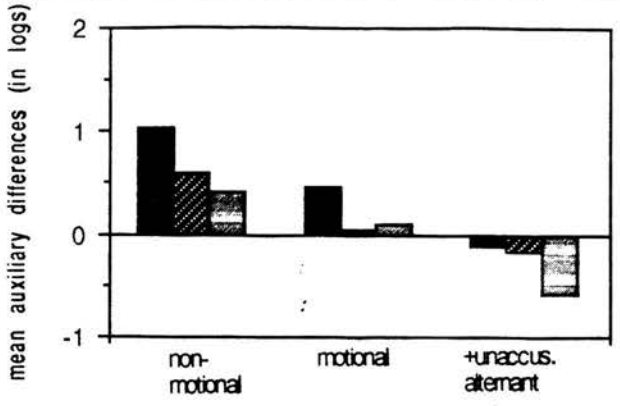
The same means are visualized in Figures 9.2a, b:

FIGURE 9.2: Unergative verbs, mean auxiliary preferences (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers)

(a)



(b) UNERGATIVE, NE-CLITICIZATION / AUXILIARY PREFERENCES



The graphs in Fig. 9.2a for basic sentences clearly show the similarity of judgment among the three language groups. All groups distinguish between non-motional and motional unergative verbs, on the one hand, and [+unaccusative alternant] on the other. The first two verb categories elicit the same strong preferences for AVERE, whereas the latter category is associated with weaker auxiliary preferences, still in favour of AVERE.

There is a highly significant main effect of semantic category, F_1 (2, 154) = 61.47, p <.0001, F_2 (2, 18) = 28.24, p <.0001, Min F' (2, 37) = 19.35, p <.01, which does not interact with native language. Further, Tukey tests yield significant differences - for all groups - between the much stronger preferences associated with non-motional verbs in basic sentences and the weaker ones for +unaccusative verbs (English: q = 32.89; French: q = 32.23; Italians: q = 29.87); also, between the stronger preferences for motional verbs and the weaker ones for [+unaccusative alternant] verbs in basic sentences (English: q = 25.1; French: q = 27.3; Italian: q = 29.55). All groups have significantly larger auxiliary differences for non-motional verbs and +unaccusative verbs in Ne-cliticization sentences (English: q = 29.27; French: q = 18.91; Italian: q = 25.01). No group has any significant difference between non-motional and motional verbs.

Finally, there are no significant differences among subject groups with respect to any verb category. The prediction is therefore borne out that informants, regardless of their native language, discriminate among unergative verbs in the direction of the hierarchy posited.

Fig. 9.2b indicates a different picture for Ne-cliticization sentences, as also shown by a significant main effect of word order, F_1 (1, 77) = 91.60, p <.0001, F_2(1, 18) = 43.77, p <.0001, Min F' (1, 37) = 29.62, p <.01.

Here, the three groups agree with respect to the direction of preferences, which favour AVERE with non-motional and motional verbs, and ESSERE with [+unaccusative alternant]

verbs. There is in fact a significant interaction of word order and semantic category, $F_1(17, 154) = 9.44$, $p < .0001$.

The three groups, however, differ with respect to the size of auxiliary differences. As predicted, native speakers' differences are closer to zero than the other two language groups, suggesting that they have knowledge of the unacceptability of Ne-cliticization with unergative verbs; contrary to the prediction, semantic categories influence the size of differences. The auxiliary differences in the French near-native judgments approximate those of native Italians with non-motional and motional verbs, but not with [+unaccusative alternant] verbs. The auxiliary differences of English near-native speakers are obviously more marked than those of the other two groups with non-motional and motional verbs, but not with [+unaccusative alternant] verbs. These differences among language groups are reflected by a significant interaction of word order and native language, $F_1(2, 77) = 9.45$, $p < .0002$, $F_2(2, 36) = 11.56$, $p < .0001$, Min F' (2, 104) = 5.2, $p < .01$.

The Tukey tests of significance conducted on the comparisons between basic and Ne-cliticization word orders again suggest that native Italians and French near-natives have similar knowledge of auxiliary selection, but English near-native speakers diverge from them. While native speakers have significantly stronger auxiliary preferences for basic sentences than for Ne-cliticization sentences with all categories of unergative verbs (non-motional: $q = 28.79$; motional: $q = 36.37$; [+unaccusative alternant]: $q = 23.93$), French near-native speakers discriminate between word orders only with non-motional ($q = 25.6$) and motional verbs ($q = 34.26$), but not with [+unaccusative alternant] verbs. English near-native speakers, on the other hand, do not differentiate between word orders with any verb category.

9.1.3 Summary of findings

The analysis of the judgments provided by native and near-native informants on unergative verbs largely supports the prediction stated in section 9.1.

All subjects differentiate among types of unergative verbs. The clearest distinction in the responses of native speakers is that between non-motional verbs, which are perceived as the ones most strongly associated with the characteristics of unergativity (selection of AVERE in the present perfect and impossibility of Ne-cliticization), and [+unaccusative alternant] verbs, which are only loosely associated with such characteristics (and therefore more acceptable when conjugated with ESSERE, or when undergoing Ne-cliticization). Thus, in contrast to the prediction that lexical-semantic distinctions would affect only judgments on basic sentences, the evidence indicates that Ne-cliticization is not a purely syntactic phenomenon. Despite the theoretical inadmissibility of either auxiliary, Ne-

cliticization with ESSERE is not uniformly rejected, but depends on the type of unergative verb: least acceptable with non-motional verbs, and most acceptable with [+unaccusative alternant] verbs. The fact that native judgments on Ne-cliticization with AVERE do not reproduce the same pattern confirms the strong relationships between Ne-cliticization and ESSERE, and between unergative verbs and AVERE.

The results also confirm the prediction that French near-natives and Italian natives judge unergativity and auxiliary selection in similar ways, but English near-natives give different judgments. The main difference is that English near-natives, unlike the other two groups, tend not to differentiate between basic and Ne-cliticized sentences.

9.2 Unaccusative verbs

Let us now turn to the pattern of acceptability found for unaccusative verbs.The predictions, for this class of verbs, were that:

(a) both native speakers and near-native speakers would differentiate among types of unaccusative verbs, so that most determinate judgments on auxiliary selection would be given to verbs on the left and least determinate judgments to verbs on the right of the following hierarchy:

change-of-location > continuation-of-state > existence-of-state > [+transitive alternant] > [+unergative alternant]

(b) native speakers would accept both basic and Ne-cliticized sentences with unaccusative verbs, but their judgments on Ne-cliticization sentences would not vary according to verb category because of the hypothesized insensitivity of Ne-cliticization to lexical-semantic representations.

(c) French and English near-native speakers would differ in their judgments on Ne-cliticization sentences, with French near-native judgments being more similar than English near-native judgments to the Italian acceptability values.

The following statistical treatments were applied to the data:

(a) a pair of four-way repeated-measures ANOVAs (proficiency level x semantic category x word order x auxiliary);

(b) Min F' values, computed on pairs of significant F_1 and F_2 values;

(c) Tukey tests on post-hoc comparisons between the F_1 means. The critical value of q at $p<.01$ for all comparisons is 6.22.

9.2.1 All variables

The geometric means of the acceptability judgments on unaccusative verbs are given in Table 9.3:

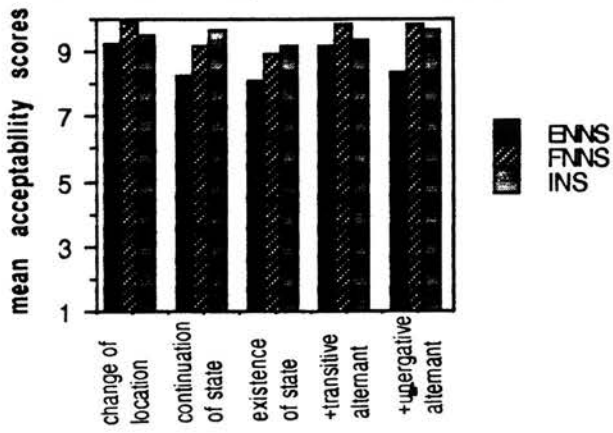
TABLE 9.3: Mean acceptability judgments on unaccusative verbs (ENNS = English near-native speakers; FNNS = French near-native speakers; INS = native Italian speakers)

	ENNS	FNNS	INS
CHANGE-OF-LOCATION, BASIC, ESSERE	9.248	9.927	9.509
CHANGE-OF-LOCATION, BASIC, AVERE	2.224	1.875	1.653
CHANGE-OF-LOCATION, NE-CLITICIZATION, ESSERE	9.323	9.109	8.540
CHANGE-OF-LOCATION, NE-CLITICIZATION, AVERE	2.038	1.709	1.697
CONTINUATION-OF-STATE, BASIC, ESSERE	8.228	9.160	9.704
CONTINUATION-OF-STATE, BASIC, AVERE	3.214	3.417	2.665
CONTINUATION-OF-STATE, NE-CLITICIZATION, ESSERE	8.332	7.874	8.465
CONTINUATION-OF-STATE, NE-CLITICIZATION, AVERE	3.017	2.085	1.975
EXISTENCE-OF-STATE, BASIC, ESSERE	8.088	8.930	9.203
EXISTENCE-OF-STATE, BASIC, AVERE	3.930	6.375	2.922
EXISTENCE-OF-STATE, NE-CLITICIZATION, ESSERE	6.874	6.801	7.198
EXISTENCE-OF-STATE, NE-CLITICIZATION, AVERE	3.914	2.892	2.667
[+TRANSITIVE ALTERNANT], BASIC, ESSERE	9.170	9.825	9.340
[+TRANSITIVE ALTERNANT], BASIC, AVERE	5.353	6.629	3.562
[+TRANSITIVE ALTERNANT], NE-CLITICIZATION, ESSERE	8.013	4.580	7.241
[+TRANSITIVE ALTERNANT], NE-CLITICIZATION, AVERE	5.798	3.810	3.911
[+UNERGATIVE ALTERNANT], BASIC, ESSERE	8.367	9.874	9.686
[+UNERGATIVE ALTERNANT], BASIC, AVERE	5.448	7.045	4.204
[+UNERGATIVE ALTERNANT], NE-CLITICIZATION, ESSERE	8.331	5.828	8.080
[+UNERGATIVE ALTERNANT], NE-CLITICIZATION, AVERE	5.344	5.083	3.715

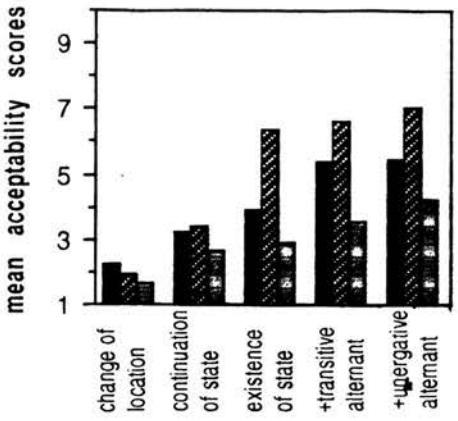
The graphic representations of the results are in Figures 9.3 a-d.

FIGURE 9.3: Mean acceptability judgments on unaccusative verbs (ENNS = English near-native speakers; FNNS = French near-native speakers; INS = native Italian speakers)

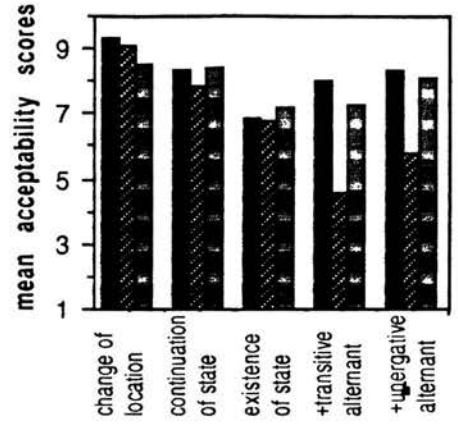
(a) **UNACCUSATIVE, BASIC / ESSERE**



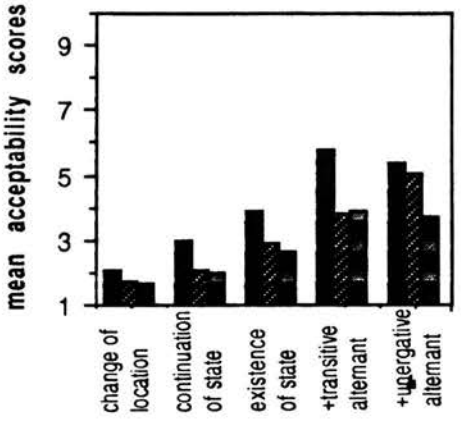
(b) UNACCUSATIVE, BASIC / *AVERE



(c) UNACCUSATIVE, NE-CLITICIZATION / ESSERE



(d) UNACCUSATIVE, NE-CLITICIZATION / *AVERE



Figures 9.3a and 9.3b, which show the pattern of responses for basic sentences with AVERE and with ESSERE, suggest that the acceptability ratings on unaccusatives with AVERE are lowest with change-of-location verbs and highest with [+unergative alternant] verbs, and that the other three categories are scaled between the extremes, consistently

with the predicted hierarchy. This pattern is basically the same for the three subject groups, although there is language-related variation with respect to the size of the individual means. The ANOVA gives a significant main effect of semantic category, F_1 (4, 308) = 32.85, p <.0001, F_2 (4, 30) = 5.34, p <.002, Min F' (4, 40) = 4.59, p <.01. Fig. 9.3b, however, indicates that sentences with ESSERE are judged as equivalent, regardless of semantic category: this gives a significant main effect of auxiliary, F_1(1, 77) = 337.06, p <.0001, F_2 (1, 30) = 227.76, p <.0001, Min F' (1, 72) = 135.72, p <.01, and a significant interaction of semantic category and auxiliary, F_1 (4, 308) = 61.16, p <.0001, F_2 (4, 30) = 11.97, p <.0001, Min F' (4, 43) = 10.01, p <.01. Finally, there is a four-way interaction of semantic category, word order, auxiliary and native language, F_1 (8, 308) = 4.77, p <.04, which qualifies all the mentioned effects.

Tukey tests on comparisons among verb categories reveal a pattern of significant differences consistent with the unaccusative hierarchy. First, there are no differences concerned with auxiliary ESSERE. With respect to sentences containing AVERE, all groups assign significantly lower acceptability ratings to change-of-location unaccusative verbs than to [+unergative alternant] verbs in both basic (English: q = 9.78; French: q = 14.44; Italian: q = 10.18) and in Ne-cliticization word order (English: q = 10.51; French: q = 11.89; Italian: q = 8.54).

All groups also judge change-of-location verbs significantly less acceptable with AVERE than [+transitive alternant] verbs, again in both basic (English: q = 9.78; French: q = 13.78; Italian: q = 8.37) and Ne-cliticization sentences (English: q = 11.4; French: q = 8.75; Italian: q = 9.1). Similarly, all groups distinguish between the lower acceptability of change-of-location verbs with AVERE and the higher acceptability of existence-of-state verbs in basic sentences (English: q = 6.42; French: q = 13.35; Italian: q = 6.41). Only English near-natives have a significant difference between these two verb categories in Ne-cliticization sentences (q = 7.11).

The three subject groups also agree in making a significant difference between lower acceptability scores given to continuation-of-state verbs and the higher scores given to [+unergative alternant] verbs in Ne-cliticization sentences (English: q = 6.24; French: q = 9.72; Italian: q = 6.89), and to [+transitive alternant] verbs (English: q = 7.13; French: q = 6.58; Italian: q = 7.45).

Figs.9.3c,d present a very similar pattern for Ne-cliticization sentences with AVERE and ESSERE. Like the judgments on basic sentences, those on Ne-cliticization with AVERE follow the predicted hierarchy. Judgments on ESSERE (see Fig.3d) are not as uniform as the equivalent ones on basic sentences, as is particularly clear in the responses of French

near-natives, but they indicate a pattern of acceptance that is not affected by the unaccusative hierarchy, or at least not to the same extent as basic sentences. Statistically, this is reflected by a significant main effect of word order, F_1 (1, 77) = 33.18, p <.0001, F_2 (1, 30) = 8.24, p <.008, Min F' (1, 46) = 6.6, p <.05, and by an interaction of native language and word order, F_1 (59, 77) = 9.49, p <.0002, F_2(59, 60) = 4.31, p <.02.
The Tukey tests on comparisons between word orders confirm the similarity of the responses on basic and Ne-cliticization sentences: no significant differences are found for any of the language groups.

The post-hoc tests of significance on the comparisons between auxiliaries suggest that, while natives and near-natives follow the same trends in their judgments, native knowledge of auxiliary selection with unaccusative verbs is more complete than near-native knowledge. Only native Italian speakers have a significant preference for ESSERE over AVERE with respect to all types of unaccusative verbs in both word orders (change-of-location, basic: q = 19.09, Ne-cliticization: q = 17.62; continuation-of-state, basic: q = 14.09, Ne-cliticization: q = 15.87; existence-of-state, basic: q = 12.51, Ne-cliticization: q = 10.83; [+transitive alternant], basic: q = 10.52, Ne-cliticization: q = 6.72; [+unergative alternant], basic: q = 9.1, Ne-cliticization: q = 8.48).
Near-native speakers exhibit a similar pattern, but they do not discriminate between auxiliaries with the two most peripheral categories of unaccusatives. English near-native speakers prefer ESSERE with change-of-location verbs (basic: q = 15.55; Ne-cliticization: q = 16.58), continuation-of-state verbs (basic: q = 10.25; Ne-cliticization: q = 11.08), and existence-of-state verbs (basic: q = 7.77).
Similarly, French near-natives give higher acceptability ratings to ESSERE in change-of-location verbs (basic: q = 18.18; Ne-cliticization: q = 18.25), continuation-of-state verbs (basic: q = 10.75; Ne-cliticization: q = 14.49), and existence-of-state verbs (Ne-cliticization: q = 9.33).
Further evidence for the language-related differences with respect to the choice of auxiliaries is provided in the ANOVA by a significant interaction of auxiliary and native language, F_1 (2, 77) = 6.7, p <.002, F_2 (2, 60) = 9.57, p <.0002.

9.2.2 Auxiliary preferences

Knowledge of auxiliary assignment with unaccusative verbs was tested more directly by analyzing the differences obtained from subtracting AVERE from ESSERE judgments. It was predicted that:

(a) all subject groups - irrespective of native language - would have stronger preferences for ESSERE, and therefore larger auxiliary differences, which range from a maximum in basic sentences for change-of-location verbs though the other categories to an indeterminate minimum for [+unergative alternant] verbs;

(b) native speakers would have auxiliary differences of the same size for unaccusative verbs in Ne-cliticization sentences. French near-native speakers would have a pattern of auxiliary preferences for Ne-cliticization sentences that approximates the Italian pattern more than the English pattern.

In order to test these predictions, the following analyses were carried out:

(a) a pair of three-way repeated-measures ANOVAs (proficiency level x semantic category x word order);

(b) Min F' values calculated on pairs of significant F_1 and F_2 values;

(c) post-hoc Tukey tests of significance applied on the F_1 means. The critical value of q for all comparisons is 8.9.

The mean auxiliary differences (in logarithmic form) obtained for the three language groups are reported in Table 9.4. As in the previous analyses, positive numbers express a preference for ESSERE, negative numbers express a preference for AVERE, and the magnitude of numbers is proportional to the strength of preferences.

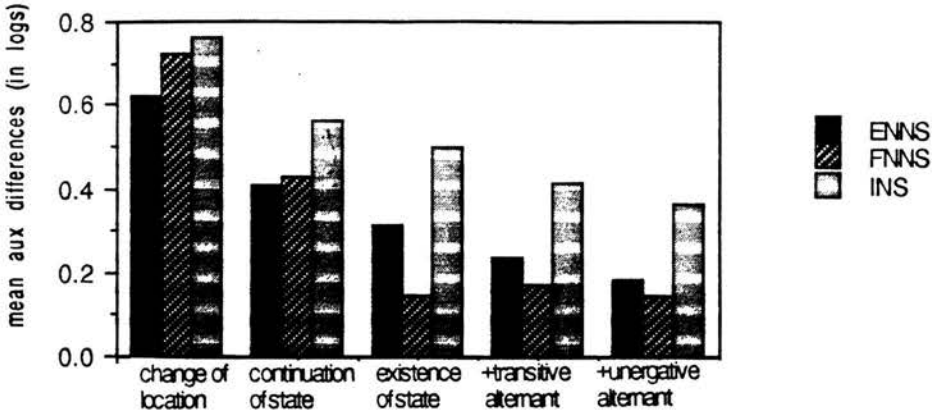
TABLE 9.4: Unaccusative verbs, mean auxiliary preferences (ESSERE - AVERE) (ENNS = English near-native speakers; FNNS = French near-native speakers; INS = native Italian speakers)

	ENNS	FNNS	INS
CHANGE-OF-LOCATION, BASIC	0.619	0.724	0.760
CHANGE-OF-LOCATION, NE-CLITICIZATION	0.660	0.727	0.702
CONTINUATION-OF-STATE, BASIC	0.408	0.428	0.561
CONTINUATION-OF-STATE, NE-CLITICIZATION	0.441	0.577	0.632
EXISTENCE-OF-STATE, BASIC	0.313	0.146	0.498
EXISTENCE-OF-STATE, NE-CLITICIZATION	0.245	0.371	0.431
[+TRANSITIVE ALTERNANT], BASIC	0.234	0.171	0.419
[+TRANSITIVE ALTERNANT], NE-CLITICIZATION	0.141	0.080	0.268
[+UNERGATIVE ALTERNANT], BASIC	0.186	0.147	0.362
[+UNERGATIVE ALTERNANT], NE-CLITICIZATION	0.193	0.059	0.337

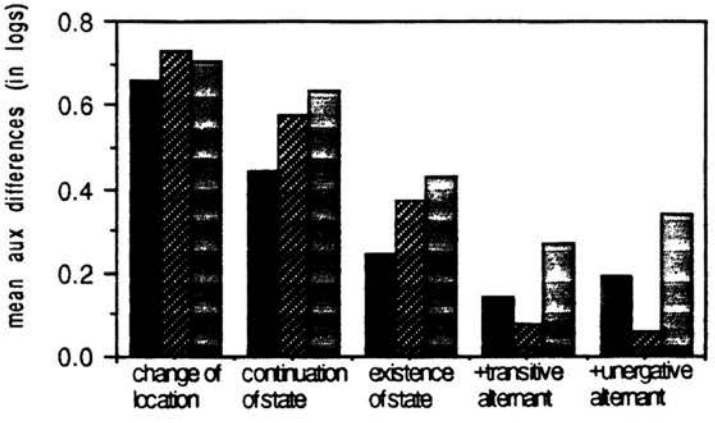
The data are also represented graphically in Figures 9.4a, b.

FIGURE 9.4: Unaccusative verbs, mean auxiliary preferences (ENNS = English near-native speakers; FNNS = French near-native speakers; INS = native Italian speakers)

(a) UNACCUSATIVE, BASIC / AUXILIARY PREFERENCES



(b) UNACCUSATIVE, NE-CLITICIZATION / AUXILIARY PREFERENCES



It is clear from Figures 9.4a and 9.4b that the patterns of judgments for basic and for Ne-cliticization sentences are very similar. The order and distribution of auxiliary preferences in both word orders follows the predictions related to the unaccusative hierarchy, with the main effect of semantic category, F_1 (4, 308) = 61.16, $p < .0001$, F_2 (4, 30) = 11.97, $p < .0001$, Min F' (4, 43) = 10.01, $p < .01$.

Tukey tests of significance on comparisons between preferences in basic and Ne-cliticization sentences do not give any significant difference, thus confirming the equivalence of the two word orders in the informants' judgments.

It is also apparent from both figures that the judgments of the three subject groups are not identical. While native Italians have stronger preferences for ESSERE, and therefore larger differences, in all cases (with the marginal exception of change-of-location verbs in Ne-cliticization word order), French near-native speakers' preferences, contrary to prediction, are not always closer to native Italians' than those of English near-native speakers. They seem to be native-like with respect to topmost core verb categories, but not with respect to peripheral categories. These language-related differences are reflected by a significant main effect of native language, F_1 (2,77) = 6.7, $p < .002$, $F_2(2, 60) = 9.57$, $p < .0002$, Min F' (2, 136) = 3.94, $p < .05$. The lack of any interactions with semantic category, or with word order, confirms that the differences among language lie in the relative size of preferences, not in the ordering of such preferences, or in any discrepancy between basic and Ne-cliticization word orders. There is, however, a marginally significant interaction of semantic category, word order and native language, F_1 (8, 308) = 4.77, $p < .04$.

The pattern revealed by the Tukey tests on comparisons among verb categories is consistent with the trends just described. All groups have stronger preference for ESSERE in basic sentences with change-of-location verbs than with [+transitive alternant] verbs (English: q = 9.67; French: q = 13.88; Italian: q = 8.96), and [+unergative alternant] verbs (English: q = 10.86; French: q = 14.49; Italian: q = 9.97). French near-natives have a larger auxiliary difference for change-of-location unaccusatives than for existence-of-state unaccusative in basic sentences (q = 14.5).
There is even more complete agreement among the three language groups with respect to Ne-cliticization sentences: all informants have stronger preferences for ESSERE with change-of-location verbs than with existence-of-state verbs (English: q = 10.43; French: q = 8.99; Italian: q = 9.2), [+transitive alternant] verbs (English: q = 13.04; French: q = 16.23; Italian: q = 10.91), and [+unergative alternant] verbs (English: q = 11.73; French: q = 16.75; Italian: q = 9.15).

9.2.3 Summary of findings

The analysis of the judgments obtained on unaccusative verbs is mostly consistent with our predictions. Both native and non-native acceptability judgments are consistent with the Unaccusative Hierarchy. Change-of-location verbs appear to be the most prototypical category of unaccusative verb, in the sense that they are the most strongly related to the selection of ESSERE. At the opposite extreme, [+unergative alternant] verbs are shown to be the least prototypical unaccusative verbs, because they are judged as the most acceptable with AVERE. The other unaccusative verb types lie between the extremes in the

predicted order, and are perceived as more or less clearly associated with ESSERE depending on their position along the proposed hierarchy.

What the predictions had not anticipated was that the effects of the unaccusative hierarchy would be manifested not only in basic sentences, but also in Ne-cliticized ones. Yet, the pattern of judgments on Ne-cliticization with AVERE closely resembles the pattern obtained for basic sentences with the same auxiliary. As in the case of unergative verbs, this strongly suggests a link between the syntactic and the semantic properties of unaccusativity.

Near-native groups did not, however, differ from native Italians in distinct ways. The main boundary separates native from near-native intuitions: although the three subject groups broadly follow the same trend in their judgments, native speakers are more discriminating than near-natives.

9.3 Restructuring verbs

The final part of this chapter is concerned with the acceptability judgments given by native and near-native speakers on restructuring constructions. The prediction was that:

(a) native speakers would accept both AVERE and ESSERE in basic and no-clitic-movement sentences, but they would accept only AVERE in clefted sentences and only ESSERE with clitic-climbing (this is because syntactic restructuring, and the AVERE --> ESSERE auxiliary change as its visible manifestation, are optional with basic and no-clitic-movement, impossible with clefting and obligatory with clitic-climbing).

(b) native speakers would not make any distinctions among type of verb inducing restructuring (i.e. Raising or Control).

(c) near-native speakers would approximate native intuitions about auxiliary choice with at least some restructuring constructions, but French near-native speakers would do so more successfully or completely than English near-native speakers.

In order to test these hypotheses, the following statistical analyses were carried out:

(a) a pair of four-way repeated-measures ANOVAs (proficiency level x verb type (raising vs control) x syntactic category (restructured vs non-restructured) x word order);

(b) Min F' values on all pairs of significant F_1 and F_2 values;

(c) post-hoc Tukey tests of significance on differences between F_1 means. The critical value of q for all comparisons is 6.86.

Notice that the distinction between AVERE and ESSERE corresponds to the distinction between 'non-restructured' and 'restructured', and will be used in place of the latter.

9.3.1 All variables

The geometric means of the acceptability judgments obtained on restructuring constructions are reported in Table 9.5.

TABLE 9.5: Mean acceptability judgments on Restructuring constructions (ENNS = English near-native speakers; FNNS = French near-native speakers; INS = native Italian speakers)

	ENNS	FNNS	INS
RAISING, BASIC, NON-RESTRUCTURED	6.977	9.420	9.749
RAISING, CLEFTING, NON-RESTRUCTURED	3.644	4.637	5.817
RAISING, CLITIC-CLIMBING, NON-RESTRUCTURED	6.623	4.285	3.143
RAISING, NO-CLITIC-MOVEMENT, NON-RESTRUCTURED	6.211	7.841	8.779
RAISING, BASIC, RESTRUCTURED	7.232	3.824	9.260
RAISING, CLEFTING, RESTRUCTURED	3.075	3.493	3.039
RAISING, CLITIC-CLIMBING, RESTRUCTURED	6.286	8.521	8.587
RAISING, NO-CLITIC-MOVEMENT, RESTRUCTURED	6.784	3.984	8.159
CONTROL, BASIC, NON-RESTRUCTURED	6.816	10.714	9.265
CONTROL, CLEFTING, NON-RESTRUCTURED	3.090	5.288	5.466
CONTROL, CLITIC-CLIMBING, NON-RESTRUCTURED	7.704	3.156	3.443
CONTROL, NO-CLITIC-MOVEMENT, NON-RESTRUCTURED	7.610	8.599	8.279
CONTROL, BASIC, RESTRUCTURED	5.876	4.024	7.081
CONTROL, CLEFTING, RESTRUCTURED	2.974	2.668	3.062
CONTROL, CLITIC-CLIMBING, RESTRUCTURED	6.259	5.838	8.768
CONTROL, NO-CLITIC-MOVEMENT, RESTRUCTURED	6.028	3.170	8.294

The data are visualized in Fig. 9.5 (Raising verbs) and Fig. 9.6 (Control verbs)

FIGURE 9.5: Mean acceptability judgments on Restructuring (Raising) verbs (ENNS = English near-native speakers; FNNS = French near-native speakers; INS = native Italian speakers)

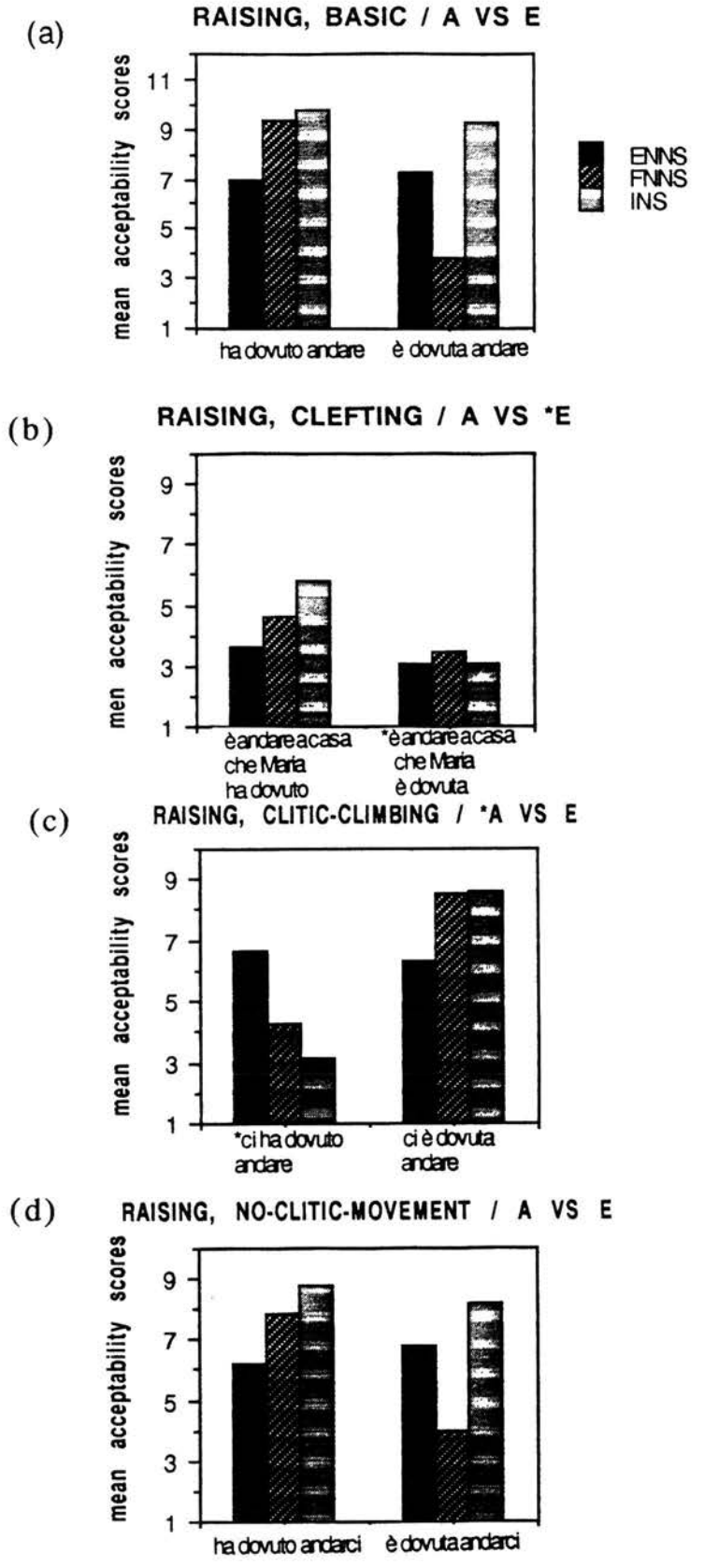
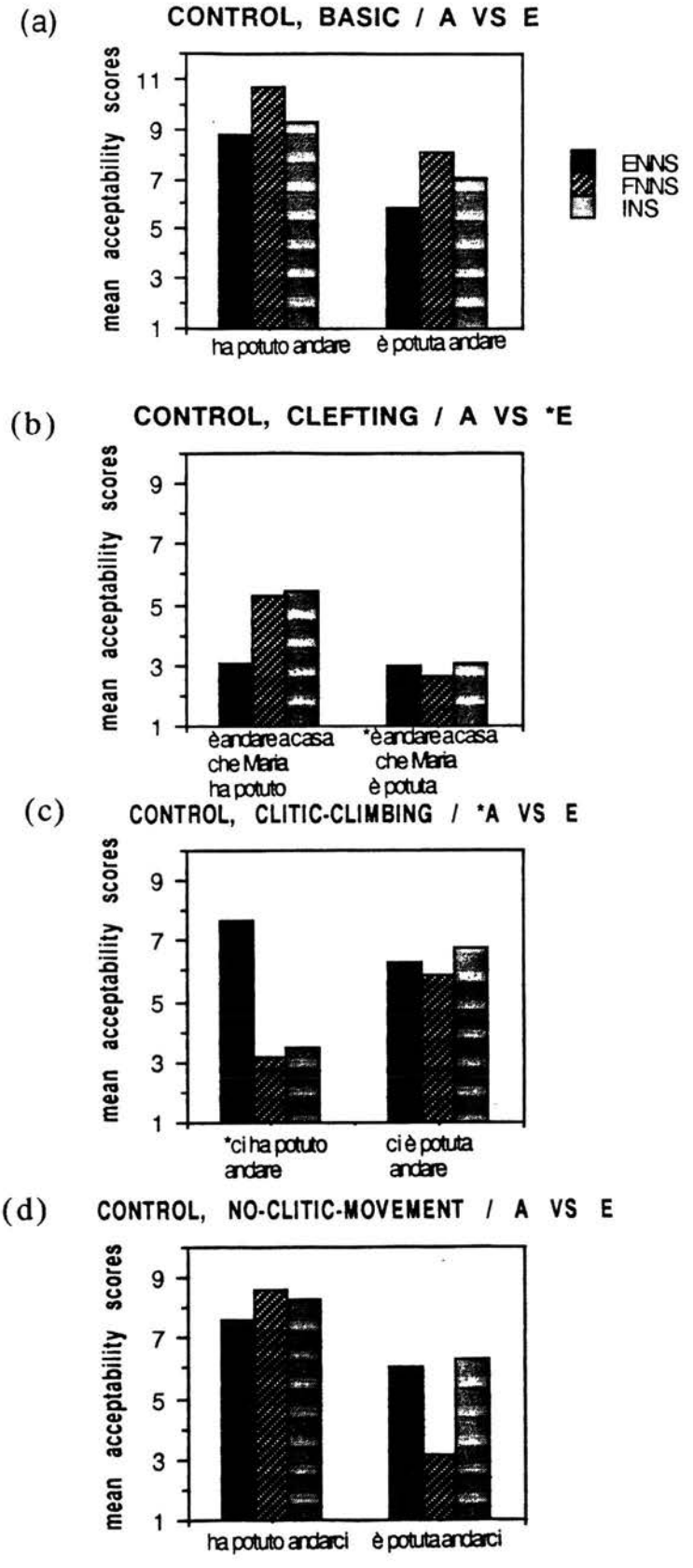


FIGURE 9.6: Mean acceptability judgments on Restructuring (Control) verbs (ENNS = English near-native speakers; FNNS = French near-native speakers; INS = native Italian speakers)



The most striking features of Figures 9.5 and 9.6 are the dissimilarities among the three language groups for restructured sentences in basic (Figures 9.5a and 9.6a) and no-clitic-movement (Figures 9.5d and 9.6d), as well as to non-restructured sentences exhibiting clitic-climbing (Figures 9.5c and 9.6c). Native speakers' judgments are determinate in accepting correct auxiliaries and in rejecting "incorrect" ones; English near-natives are indeterminate with respect to both optional and obligatory auxiliaries; French near-natives are categorical (i.e. accepting only one auxiliary even when both are allowed).

There is strong statistical evidence supporting this interpretation. Given that ESSERE is optional with some word orders and obligatory with others, it is not surprising to find that syntactic category and word order interact, $F_1(3, 231) = 36.69$, $p < .0001$, $F_2(3, 24) = 11.77$, $p < .0001$, Min F' $(3, 41) = 8.91$, $p < .01$. The obvious language-related differences among informants determine an interaction of syntactic category and native language, $F_1(2, 77) = 8.36$, $p < .0005$, $F_2(2, 48) = 5.73$, $p < .006$, Min F' $(2, 105) = 3.4$, $p < .05$, an interaction of word order and native language, $F_1(6, 231) = 5.56$, $p < .0001$, $F_2(6, 48) = 5.57$, $p < .0002$, Min F' $(6, 161) = 2.78$, $p < .05$, a three-way interaction of syntactic category, word order and native language, $F_1(6, 231) = 11.71$, $p < .0001$, $F_2(6, 48) = 9.41$, $p < .0001$, Min F' $(6, 138) = 5.22$, $p < .01$, and a four-way interaction of syntactic category, word order, native language and verb type, $F_1(6, 231) = 2.84$, $p < .01$.

Let us now look more closely at the differences among native languages. For non-restructured sentences (i.e. those with AVERE) Tukey tests indicate that English near-native speakers give significantly higher acceptability ratings than native Italian speakers in sentences with clitic-climbing (raising: $q = 8.13$; control: $q = 8.79$). They also give higher acceptability ratings than French near-natives on the same kind of construction, but only with control verbs ($q = 9.74$).

For restructured constructions (i.e. those with ESSERE), the tests reveal significant differences between the low acceptability scores given by French near-natives to basic sentences and the higher scores given by Italians (raising: $q = 9.64$; control: $q = 6.87$) and by English near-natives (raising: $q = 6.95$). The same pattern is found for the other construction that allows optional restructuring, that is, no-clitic movement: French near-natives assign significantly lower acceptability ratings to these sentences with ESSERE than native Italians do (raising: $q = 7.82$; control: $q = 10.49$). French near-natives seem therefore not to recognize the optionality of restructuring with basic and no-clitic-movement sentences: in their linguistic intuitions, only AVERE is the correct auxiliary.

The other major differences between natives and near-natives concerns clitic-climbing in the non-restructured syntactic condition. Unlike the other two groups, English near-native speakers tend to accept clitic-climbing with AVERE, as shown by the significant differences between their acceptability ratings and those of Italians (raising: q = 8.13; control: q = 8.79), and of French near-natives (control: q = 9.74).

The Tukey tests on comparisons between the non-restructured and the restructured syntactic categories provide the most telling evidence about the ability of the three language groups to discriminate between auxiliaries. The judgments of native Italian speakers conform to the prediction, in that there are significant differences between the two syntactic conditions for the constructions that allow only one auxiliary, whereas there are no differences for the constructions that allow both. Thus, native speakers assign higher acceptability ratings to AVERE than to ESSERE in clefted sentences (raising: q = 7.08; control: q = 6.92), but prefer ESSERE to AVERE with clitic-climbing (raising: q = 10.96; control: q = 10.2). They do not make any significant distinctions between auxiliaries for basic and no-clitic-movement word orders.

The judgments of French near-natives, on the other hand, discriminate between auxiliaries in almost all cases. They judge AVERE more acceptable than ESSERE in basic sentences (raising: q = 9.83; control: q = 10.68), in no-clitic-movement sentences (raising: q = 7.39; control: q = 10.88), and in clefted sentences, but only with control verbs (q = 7.46); however, they prefer ESSERE to AVERE with clitic-climbing (raising: q = 7.49; control: q = 6.91). This confirms that the competence of French near-natives is as determinate as the competence of native speakers, but includes a representation of auxiliary selection which diverges substantially from the native representation.
Finally, there are no significant differences between auxiliaries in the responses of English near-natives: their judgments are indeterminate.

Tukey tests conducted on the comparisons among word orders within language groups again indicate that native Italians are the only group whose knowledge of auxiliary selection with restructuring corresponds to the predictions. For non-restructured (i.e. AVERE) sentences, their acceptability ratings for clitic-climbing are significantly lower than those for basic sentences (raising: q = 12.34; control: q = 10.8), clefting (only with raising verbs: q = 6.91), and no-clitic-movement (raising:q = 11; control: q = 9.57). This is consistent with the predictions, since clitic-climbing is the only construction that does not allow AVERE.
For restructured (i.e. ESSERE) sentences, native acceptability scores on clefting are significantly lower than those on basic word order (raising: q = 12.15; control: q = 9.15), clitic-

climbing (raising: q = 11.33; control: q = 11.48), and no-clitic-movement (raising: q = 10.77; control: q = 10.87). This is also predictable, because clefting is the only construction that does not allow ESSERE.

The same analysis on the French near-natives' judgments reveals that, in non-restructured sentences, these informants assign lower acceptability ratings to clitic-climbing than to basic word order (raising: q = 8.59; control: q = 13.33) and no-clitic movement (but only with control verbs: q = 10.93). Unlike native speakers, they also have significantly lower acceptability scores for clefted than for basic sentences (raising: q = 7.73; control: q = 7.7). For restructured (i.e.ESSERE) sentences, French near-natives have a significant preference for clitic-climbing over basic (raising: q = 8.73), clefting (raising: q = 9.72; control: q = 8.54), and no-clitic-movement (raising: q = 8.29; control: q = 6.96).
English near-natives have significantly lower acceptability ratings for clefted sentences in the non-restructured condition than for basic (raising: q = 7.08; control: q = 8.63), clitic-climbing (control: q = 9.99), and no-clitic-movement (control: q = 9.82). They have the same pattern of preferences for restructured sentences: clefting is judged significantly less acceotable than basic (raising: q = 9.38; control: q = 7.42), clitic-climbing (raising: q = 7.8; control: q = 8.11), and no-clitic-movement (raising:q = 8.63; control: q = 7.7). Given the relative markedness of clefted sentences, and the absence of other significant differences, this pattern is not indicative of grammatical competence: it is plausible to draw the conclusion that English near-native speakers have indeterminate knowledge of auxiliary selection under restructuring.

The results analysed so far have shown that raising and control verbs induce roughly the same type of response, as can be seen from the shape of the graphs in Figs. 9.5 and 9.6. There is no significant main effect of verb type, and the only significant interaction is the already mentioned four-way one of syntactic category, word order, native language and verb type, F_1 (6, 231) = 2.84, $p < .01$. Tukey tests do not give any significant differences between the two verb types for any word order and at any level. On the basis of these results, the two verb types seem therefore equivalent with respect to syntactic restructuring.

9.3.2 Auxiliary preferences

As was done for the other verb classes, mean auxiliary differences were calculated by subtracting the acceptability scores for one auxiliary from the scores for the other. Given the different requirements of the four word orders investigated with respect to auxiliary selection, the differences were computed as follows:

- BASIC: AVERE - ESSERE (expected difference = 0)
- CLEFTING: AVERE - ESSERE (expected difference > 0)
- CLITIC-CLIMBING: ESSERE - AVERE (expected difference > 0)
- NO CLITIC MOVEMENT: AVERE - ESSERE (expected difference = 0)

It was predicted that:

(a) native speakers would have no auxiliary preferences for basic and no-clitic-movement sentences (and therefore differences close to zero), but they would have a preference for ESSERE with clitic-climbing and a preference for AVERE with clefting (resulting in large differences)

(b) neither French near-native speakers, nor English near-native speakers would share a similar pattern of auxiliary preferences with native Italians, but the French informants would show a closer approximation to native speakers than the English.

(c) informants would give similar responses for raising and control verbs.

These hypotheses were tested by means of:

(a) a pair of three-way repeated measures ANOVAs (proficiency level x verb type x word order);

(b) Min F' values for pairs of significant F_1 and F_2 values;

(c) post-hoc Tukey tests of significance were applied to selected comparisons between F_1 means. The critical value of q at $p < .01$ for all comparisons is 8.9.

The means of the auxiliary differences (in logarithmic form) obtained are reported in Table 9.6. Positive numbers indicate a preference for AVERE, except in the case of clitic-climbing, where they indicate a preference for ESSERE. Negative numbers indicate a preference for ESSERE, except for clitic-climbing, where they indicate a preference for AVERE. Larger numbers indicate stronger preferences.

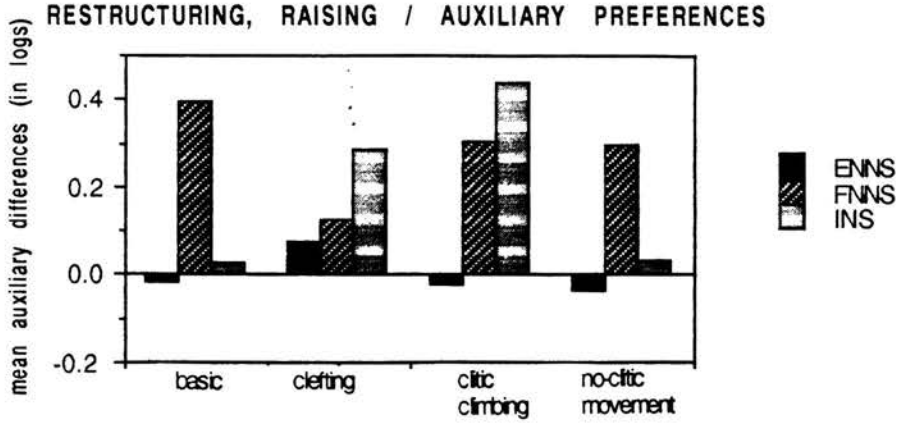
TABLE 9.6: Restructuring verbs, mean auxiliary preferences (ENNS = English near-native speakers; FNNS = French near-native speakers; INS = native Italian speakers)

	ENNS	FNNS	INS
RAISING, BASIC	-0.016	0.391	0.022
RAISING, CLEFTING	0.074	0.123	0.282
RAISING, CLITIC-CLIMBING	-0.023	0.298	0.434
RAISING, NO-CLITIC-MOVEMENT	-0.038	0.294	0.032
CONTROL, BASIC	0.064	0.425	0.117
CONTROL, CLEFTING	0.017	0.297	0.252
CONTROL, CLITIC-CLIMBING	-0.090	0.267	0.406
CONTROL, NO-CLITIC-MOVEMENT	0.101	0.433	0.001

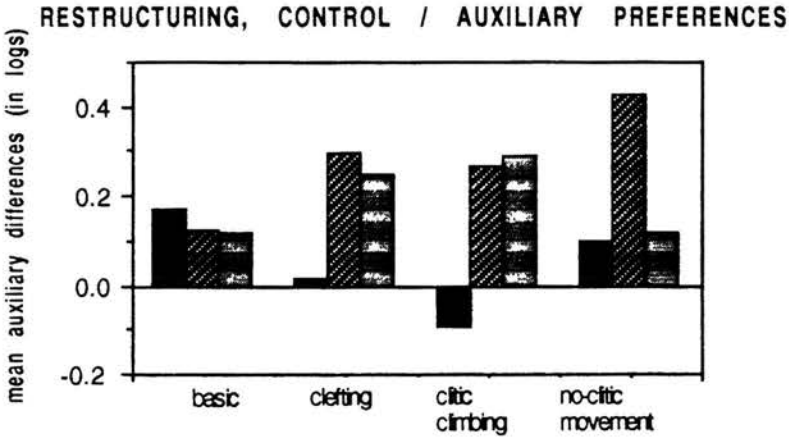
The graphic representation of the mean differences is shown in Figures 9.6a, b.

FIGURE 9.6: Restructuring verbs, mean auxiliary preferences (ENNS = English near-native speakers; FNNS = French near-native speakers; INS = native Italian speakers)

(a)



(b)



The graphs in Fig. 9.7 immediately confirm the dissimilarity of the responses given by the three language groups. While the auxiliary preferences of native Italian informants - consistent with the predictions - are marked for clefting and clitic-climbing and minimal for basic and no-clitic-movement, those of English near-native speakers are all close to zero, and those of French near-native speakers are all relatively large. This suggests that English near-natives have no clear auxiliary preferences for any of the constructions tested, and French near-natives have marked auxiliary preferences for all these constructions, independently of the optionality or obligatoriness of auxiliary assignment.

Statistically, these three different acceptability patterns are reflected in a main effect of native language, F_1 (2, 77) = 14.91, p <.0001, F_2(2, 48) = 20.24, p <.0001, Min F' (2, 124) = 8.59, p <.01 and, more crucially, by a significant interaction of native language and word order, F_1(6, 231) = 8.21, p <.0001, F_2 (6, 48) = 4.57, p <.001, Min F' (6, 110) = 2.94, p <.05.

There are no significant differences between preferences expressed with raising verbs and those expressed with control verbs. The slightly more marked preferences accorded by some informants to AVERE with control verbs, however, are reflected in a statistical interaction of verb type, word order and native language, F_1 (6, 231) = 2,75, p <.02.

Tukey tests on the comparisons among language groups also indicate that English near-natives under-discriminate between auxiliaries (i.e. they do not differentiate between correct and incorrect auxiliary choice, even when this is required), whereas French near-natives over-discriminate (i.e. they assume that only one auxiliary is correct, even when both are). This is evident from the fact that English near-native have significantly weaker auxiliary preferences for clitic-climbing than Italians (raising: q = 11.54; control: q = 12.47) and French near-natives (raising: q = 9.07; control: q = 8.98). Moreover, French near-natives have significantly stronger auxiliary preferences for basic sentences than native Italians (only for raising verbs: q = 9.28), and than English near-natives (raising: q = 10.23; control: q = 9.07); similarly, they have a significantly stronger preference for AVERE with no-clitic-movement sentences than both native speakers (raising: q = 8.9; control: q = 10.82) and English near-natives do (raising: q = 8.95; control: q = 8.91). The auxiliary preferences of the three groups are not significantly different with respect to clefting, although the means for English near-natives are noticeably lower than those of the other two groups (see Figures 9.7a, b). This suggests, as the previous results have already done, the relative markedness of clefting compared to the other constructions.

9.3.3 A comparison with the Card-Sorting results

If the results just mentioned - which were obtained via the Magnitude Estimation method - are compared with the corresponding results obtained via Card-Sorting, an interesting divergence emerges.

Card-Sorting does not in fact yield any significant differences among the patterns of judgments produced by the three language groups, as shown by the flat profiles of the graphs in Figs. C12 (Raising verbs) and C13 (Control verbs) in Appendix C. The ANOVAs run on all variables (see Tables B45 and B57 in Appendix B) gives no significant main effect of native language, and only one significant interaction of native language with other

variables (a three-way interaction of syntactic category, word order and native language, F_1 (6, 231) = 2.87, p <.01). Tukey tests reveal no differences among groups for either verb type. This picture is confirmed by the ANOVAs run on auxiliary differences (see Tables B46 and B58 in Appendix B). Although the by subject analysis gives a significant main effect of native language, F_1 (2, 77) = 4.02, p <.02, Tukey tests produce no significant difference among the auxiliary preferences expressed by the three language groups.

The other divergence between the results of Magnitude Estimation and those of Card-Sorting has to do with the relatively more sizable difference in Card-Sorting between the judgments on Raising and the judgments on Control verbs. The ANOVAs give no significant main effect of verb type (see Tables B45 and B57 in Appendix B) but the by subject analysis shows significant interactions of verb type and semantic category, F_1 (1, 77) = 8.95, p <.004, verb type and word order, F_1 (3, 231) = 6.56, p <.0003, and verb type, semantic category and word order, F_1 (3, 231) = 10.12, p <.0001. The ANOVAs on auxiliary preferences (see Tables B46 and B58) give a significant main effect of verb type, F_1 (1, 77) =25.04, p <.0001, F_2 (1, 24) = 9.85, p <.005, Min F' (1, 44) = 7.07, p <.05. These effects seem to be due to the acceptability ratings given to Control verbs with AVERE, which are higher than those given to the corresponding Raising verbs. Tukey tests on comparisons between verb types, however, do not produce any significant differences.

9.3.4 Summary of findings

The results obtained for the restructuring constructions allow us to draw two conclusions. The first is that near-native speakers have different knowledge representations of auxiliary selection from those of native speakers, despite their native-like performance. The second is that French and English near-native speakers have different views on Italian auxiliary selection.

Consistent with the predictions, native speakers are the only group who display knowledge of the subtle properties of restructuring constructions with respect to auxiliary selection, and thus the only informants who produce determinate judgments consistent with the predictions. In contrast with the predictions, however, the two near-native groups are both different from natives, but in two distinct ways. French near-natives also express determinate judgments, which are however only partly in agreement with native judgments: since they seem to have no knowledge of the optionality of auxiliary change with some constructions, they turn it into categorical choice by allowing only AVERE as the correct auxiliary. English near-natives, in contrast with both the other groups, do not give determinate judgments on

any restructuring construction: they simply do not discriminate between correct and incorrect, or optional and obligatory auxiliary choice.

The pattern of judgments on restructuring is therefore substantially different from those obtained for unergative and unaccusative verbs. The results provide evidence for the claim that phenomena which cannot be characterized semantically are not fully mastered even late in interlanguage development.

Interestingly, this pattern of differences among language groups is not reproduced in the judgments collected with the Card-Sorting method. In the untimed test conditions created by Card-Sorting, native and near-native speakers produce similar judgments on all restructuring constructions, and the two near-native groups perform similarly to each other.

CHAPTER 10
CONCLUSIONS

10.0 Introduction

This thesis was defined in the Preface as an experimental approach to systematic variation in native and non-native intuitions on Italian auxiliary selection. Our whole enterprise was based on four critical assumptions: (a) the idea that linguistic acceptability is a relative, not an absolute, property of grammars; (b) the existence of indeterminacy in native and non-native grammars, and the persistence of indeterminacy in near-native grammars; (c) the existence of systematic variation in Italian auxiliary selection, and (d) the validity and reliability of magnitude estimation as a technique for the elicitation of judgments of linguistic acceptability.

We did not have any independent evidence in support of the latter two arguments, which substantiate the former two; indeed, it was to goal of this thesis to provide such evidence. Thus, the success of this study can be measured by the extent to which the experimental results can be regarded as substantial evidence that auxiliary selection is systematically variable, and that the magnitude scaling of linguistic acceptability is feasible and meaningful.

In order to evaluate the import of our results, we will refer back to the experimental hypotheses stated in Chapter 4. We will then be in a position to decide whether the null hypotheses can be rejected. The hypotheses in Chapter 4 were grouped according to three headings: psychological, methodological, and linguistic, corresponding to the theoretical concepts that underlie this study. This chapter will follow the same order.

10.1 Significance of the results from a psychological point of view

10.1.1 Existence of indeterminacy in native grammars

This study has first and foremost shown that linguistic acceptability is not an absolute, either/or dimension. Native informants do not produce uniform judgments on sentences: if they did, they would have uniformly accepted correct sentences and uniformly rejected incorrect sentences, irrespective of type of construction or verb category. Instead, they give clear and consistent judgments on some sentences, and less clear and less consistent judgments on others, which suggests that acceptability is a gradable, more-or-less dimension. The concept of acceptability hierarchy was employed in Chapter 1 to refer to the relativity of acceptability judgments: acceptability hierarchies can be portrayed as continuous

dimensions having a determinate core and an indeterminate periphery. This is precisely what was found in this study. The intuitions of native Italian speakers are characterized by predictable indeterminacy, in the sense that, given the theoretically justified acceptability hierarchy that we proposed, the consistency of their judgments is a function of the position of constructions on the hierarchy: highest for core constructions, and progressively decreasing for other constructions according to their distance from the core.

10.1.2 Persistence of indeterminacy in non-native grammars

It was also predicted that non-native intuitions would be characterized by a higher degree of indeterminacy than native intuitions. We argued that, although inderminacy is overall more marked at lower than at highest levels of proficiency, it would not necessarily disappear at the highest proficiency levels. This prediction was also borne out by the results of this study. The judgments of near-native speakers of Italian on some of the constructions tested were clearly affected by indeterminacy, unlike the corresponding judgments given by native speakers.

10.2 Significance of results from a methodological point of view

10.2.1 Viability of Magnitude Estimation

It was pointed out in Chapter 2 that, although it is possible to posit acceptability hierarchies on the basis of theoretical arguments, such arguments predict only the relative order of constructions with respect to one another but not the distance between constructions. It was also noted that the most commonly used methods for the elicitation of acceptability judgments are based either on nominal or on ordinal scales, and therefore do not provide information on the size of the intervals between values on an acceptability scale. These methods, like any other categorical measurement, have little sensitivity to degrees of acceptability and are unlikely to capture indeterminacy of judgments. Moreover, such methods deny access to parametric statistical analyses because they do not satisfy the requirement that scores be at least on an interval scale.

These questions motivated our decision to investigate the applicability of magnitude estimation to the elicitation of acceptability judgments. Magnitude estimation is known to produce reliable interval scales both in psychophysics (which is its principal domain of application) and in social psychology, but it has never been applied to the investigation of linguistic acceptability before. The present study employed both magnitude estimation and card-sorting (a method based on an ordinal level of measurement), in order to set up a

comparison between the two. Our weak hypothesis was that magnitude estimation would be as least as informative as card-sorting; our strong hypothesis was that magnitude estimation would be more discriminating, and therefore more powerful, than card-sorting.

Our results support our experimental hypotheses. First, they indicate that magnitude estimation is a viable technique: if they are properly instructed, informants can produce ratio judgments of linguistic acceptability, and do in fact use a wider scale than with card-sorting, even though the face-validity of magnitude estimation is low. This result reinforces the concept of relative acceptability. It also stresses the fundamental inadequacy of category scales of linguistic acceptability, which are unable to capture the differentiation in informants' intuitions and are therefore bound to under-represent their grammatical competence. Second, the results give full support to the weak hypothesis stated above: magnitude estimation and card-sorting yield consistent patterns of acceptability, in the sense that they never contradict each other. More crucially, the results support the strong hypothesis as well: magnitude estimation produces a more fine-grained pattern of acceptability than card-sorting, and therefore more statistically significant results; it also reveals some patterns of acceptability that are not detected by card-sorting. We can therefore reject the null hypothesis that magnitude estimation does not work with judgments of linguistic acceptability; that it produces less discriminating and less consistent results than card-sorting; that it reveals uninterpretable patterns of acceptability.

In spite of some caveats (i.e. the lack of a calibration procedure in the research design), we can conclude on these bases that the magnitude scaling of acceptability is a feasible and powerful elicitation technique: it provides more information than other methods and it permits the statistical manipulation of such information by means of parametric analyses. This study justifies an in-depth validation of the method, and opens up a potentially fruitful new direction in linguistic research.[1]

10.3 Significance of results from a linguistic point of view

The linguistic hypotheses in Chapter 4 were presented under the two headings of *determinacy* and *learnability*, which will also be used in the following sections.

10.3.1 Reality of the Unaccusative Hypothesis

Auxiliary selection in Italian was identified as an area of grammar exhibiting systematic variation. The theoretical arguments at the foundations of this study rely on the Unaccusative

Hypothesis, that is, the division of intransitive verbs into the two classes of unaccusative and unergative verbs which have different syntactic representations. This hypothesis was the original insight that sparked most of the recent research on auxiliary selection in Italian and other languages. The Government-Binding formulation of the Unaccusative Hypothesis allows us not only to capture generalizations about the distribution and the syntactic behaviour of auxiliaries ESSERE and AVERE, but also to appreciate the relationship among apparently unrelated phenomena, such as the selection of perfective auxiliaries and the syntactic properties optionally induced by 'restructuring' verbs, and to establish the bases for a typology of unaccusativity and auxiliary selection where the auxiliary systems in different languages are in parametric varation.

The results of the present study confirm that the Unaccusative Hypothesis is reflected in native judgments on perfective auxiliary selection and NE-cliticization: native informants consistently accept ESSERE and NE-cliticization with unaccusative verbs, and AVERE, but not NE-cliticization, with unergative verbs.

As we pointed out in Chapter 3, however, the Unaccusative Hypothesis - because of its exclusive syntactic orientation - does not deal adequately with variation in auxiliary selection, either cross-linguistically or language-internally. In many cases, the syntactic generalizations attained on the basis of the Unaccusative Hypothesis are qualified by semantic factors; for languages that have very few syntactic reflexes of unaccusativity (such as English), the unaccusative/unergative distinction has to be reformulated in semantic terms.

We argued in Chapter 3 that the semantic aspects of unaccusativity are best accounted for not by theories of auxiliary selection that focus exclusively on semantic characterizations, but by theories that seek to explain the interface between the syntax and the semantics of auxiliary selection, such as theories of argument structure. These theories maintain the importance of the syntactic characterization offered by the Unaccusative Hypothesis but in addition they attempt to formalize the level of lexical-conceptual structure (which specifies the meanings of verbs) and the mapping operations which translate lexical-semantic variables into argument-structure variables (which are the input to the syntactic level of representation). This study maintained that there are phenomena of auxiliary selection which have a purely syntactic origin, i.e. that are derived from the syntactic status of arguments (such as NE-cliticization and restructuring constructions), and phenomena that have a semantic as well as a syntactic origin, i.e. that are derived from the lexical-semantic status of arguments (such as the selection of perfective auxiliaries).

The specific proposal put forward in this study is intended as a contribution to the understanding of the latter kind of phenomena. It was argued that further differentiation within the classes of unaccusative verbs in lexical-semantic terms is necessary to account for (a) the path of evolution of auxiliaries in Romance languages; (b) the synchronic variation in auxiliary usage exhibited by modern Italian and French, and (c) the acquisition of auxiliaries by learners of Italian as a foreign language. Our claim was that unaccusative verbs can be placed on an acceptability hierarchy defined by the dimensions 'dynamic vs static' and 'concrete vs abstract'. the dimensions introduce distinctions within the broad concept of 'theme' (normally interpreted as 'affected entity') according to the type of process undergone by the subject of the verb. Core verbs (i.e. the most prototypically unaccusative) are unpaired unaccusative verbs denoting change of location; the least prototypically unaccusative are paired verbs with an unergative alternant. These distinctions offer a framework for the interpretation of diachronic and synchronic data and for the formulation of specific hypotheses about second language acquisition.
It was suggested that the class of unergative verbs may also be heterogeneous: an unergative hierarchy was proposed based on the notion of '+/- motional activity', such that non-motional unergative verbs would be at the core of the hierarchy. It was recognized that this hierarchy, unlike the unaccusative hierarchy, is not supported by independent diachronic or synchronic evidence.

10.3.2 Lawful indeterminacy

We hypothesized that the acceptability hierarchies proposed for unaccusative and unergative verbs would be reflected in the acceptability judgments of native Italian speakers, in the sense that verb categories closer to the core would elicit more determinate judgments,and verb types more distant from the core would elicit indeterminate judgments.

This hypothesis is confirmed by the data. The degree of determinacy of native judgments (as measured by the strength of auxiliary preferences, for example) is related to the position of verb categories along the unaccusative and unergative hierarchies. This constitutes evidence that the linguistic acceptability of auxiliary selection is systematically variable and that indeterminacy is - to some extent - predictable.
We can therefore reject the null hypothesis that judgments on auxiliary choice with unaccusative and unergative verbs are uniform and categorical, regardless of lexical-semantic distinctions within the two verb classes.

However, conclusions about the existence of acceptability hierarchies need to be qualified. While it is indisputable that the relative order of verb types on the respective hierarchies is not contradicted by the data (i.e. verbs types closer to the core are normally judged as least as acceptable as, but not as less acceptable than, verb types more distant from the core), the distances between successive verb types on the scales are not always statistically significant. Recall that it was in order to obtain interval measurements of acceptability that we employed Magnitude Estimation: such information therefore tells us that not all the lexical-semantic distinctions along the hierachies are equally salient in the linguistic intuitions of native speakers (although the distinctions are still valid descriptive and interpretative devices). Nevertheless, the fact that native judgments respect the relative order of verb types on the hierarchies is an important result, which shows that our hypotheses are in the right direction, and that these issues deserve further investigation.

10.3.3 Semantic vs syntactic aspects of auxiliary selection

It was predicted that the distinction between semantically-based and syntactically-based aspects of auxiliary selection would be reflected in differences at the level of native acceptability judgments: semantically-based phenomena would receive systematically variable judgments (as we saw earlier), while syntactically-based phenomena would not show variability (i.e. they would tend to be categorical). This prediction is only partly supported by the data.

As far as the restructuring constructions are concerned, these are indeed judged categorically by native speakers, who accept sentences containing the correct auxiliaries and reject sentences containing the incorrect auxiliary. Restructuring constructions are on the whole judged with less determinacy than perfective auxiliary selection and NE-cliticization: this was also predicted on the basis of Burzio's configurational hierarchy for ESSERE-assignment, which identifies restructuring as peripheral.

The judgments on NE-cliticization with unaccusatives (and to a lesser extent those on unergatives), however, contradict our prediction: instead of eliciting categorical judgments, NE-cliticization appears to be conditioned by the Unaccusative hierarchy in the same way as the selection of perfective auxiliaries: it is judged most acceptable with core unaccusative verb types and least acceptable with peripheral unaccusative verb types. This surprising result is quite important because it suggests that semantic factors and syntactic factors of unaccusativity are more closely connected than it has been so far assumed, and syntax is not

'blind' to semantic information. It also opens up the possibility (only sketched at the end of Chapter 3) that syntactic change may be a consequence of lexical-semantic change.

We are therefore not able to completely reject the null hypothesis that there is no difference, in terms of variation/uniformity between native judgments on semantically-based phenomena and judgments on syntactically-based phenomena. Native informants give categorical judgments on restructuring constructions, but not on NE-cliticization. Judgments on the latter are systematically variable as judgments on perfective auxiliary selection because they are also conditioned by the lexical-semantic hierarchies.

10.4 Learnability

Let us now turn to the results obtained for non-native judgments. Our two main experimental hypotheses were concerned with development and with ultimate attainment.

10.4.1 Development

Our most basic prediction about non-native judgments was that they would be on the whole less determinate than native judgments, but that overall indeterminacy would decrease with increase in proficiency. This prediction is comfortably supported by the data: regardless of verb category and type of construction, the degree of determinacy in judgments is a function of proficiency level: beginners' judgments are the most indeterminate and near-native judgments are the most determinate. We take this to be a natural reflection of low-level learners' ignorance of Italian: as learners are exposed to more input, they acquire more knowledge and clearer intuitions about what is allowed and what is not allowed by the foreign language grammar. Determinacy of judgments is first shown by judgments on correct auxiliaries, and only later does it extend to judgments on incorrect auxiliaries. This seems to indicate that knowing a language involves greater awareness of unacceptability, rather than acceptability (as was pointed out in Chapter 7).

From a developmental point of view, we predicted that the acceptability hierarchies for unaccusative and unergative verbs would be reflected in the order of appearance of clear intuitions about auxiliary selection in learners' acceptability judgments: knowledge about auxiliary selection in the present perfect would be first acquired for core verb types (i.e. learners at lower proficiency levels would produce determinate judgments about core verbs) and then gradually be acquired for the other verb types on the hierarchies, while peripheral verb types would remain indeterminate.

The results of this study support this prediction. Learners' knowledge of auxiliary selection (as measured by the determinacy of their acceptability judgments) begins with core verb types and gradually spreads to less central verb types. This indicates that core verbs are easier to acquire because they are more salient in the input and because of the cross-linguistic validity of the verb hierarchies.
We are thus able to reject the null hypothesis that there are no differences within the class of unaccusative verbs, nor within the class of unergative verbs, with respect to ease of acquisition.

We also predicted that the distinction between semantically-based phenomena and syntactically-based phenomena would be reflected by the shape of the path of development of these phenomena in the learners' intuitions: the development of knowledge about the selection of perfective auxiliary would be gradual, cumulative, and conditioned by the acceptability hierarchies, whereas the development of knowledge of NE-cliticization and the restructuring constructions would be discontinuous and uniform for all verb types. We have just seen that the first prediction is supported by the data. The second prediction, however, is only partly borne out. Contrary to our expectations, in fact, NE-cliticization follows a similar path of development as the selection of perfective auxiliary: both are conditioned by the acceptability hierarchies, as they are in native judgments. Restructuring constructions, on the other hand, never become determinate in non-native intuitions: even near-native speakers give indeterminate judgments about optional auxiliary change in basic sentences and obligatory auxiliary change under clitic-climbing. We return to this point below.

There is an interesting divergence in this respect between Magnitude Estimation and Card-Sorting. While Magnitude Estimation produces a flat developmental profile for restructuring constructions, Card-Sorting shows that the intuitions of near-native speakers are similar to those of native speakers. One plausible interpretation of such divergence (which is of course impossible to test with any accuracy) is that near-natives have only metalinguistic knowledge of restructuring, which they can retrieve in the untimed experimental conditions provided by Card-Sorting but not in the strictly timed conditions imposed by Magnitude Estimation. If this were true, then we could claim that Magnitude Estimation taps 'tacit' grammatical knowledge, as opposed to metalinguistic knowledge, and provides a more direct picture of the informant's internalized grammar.

Thus, the null hypothesis of no differences between syntactically-based and semantically-based aspects of auxiliary selection with respect to the shape of development cannot be completely rejected. On the one hand, perfective auxiliary selection clearly presents a

gradual developmental profile, consistent with the prediction. On the other hand, and contrary to the prediction, NE-cliticization with unaccusatives also tends to follow a gradual developmental path. Restructuring constructions remain indeterminate across proficiency levels, which confirms their inherent markedness.

10.4.4 Ultimate attainment

Finally, we hypothesized that the native language of learners would determine whether interlanguage indeterminacy with respect to syntactic phenomena related to auxiliary selection is 'resolved' at the near-native level. Specifically, it was predicted that French near-native judgments would be more similar to native Italian judgments than the judgments of English near-native speakers are, because French presents *partial instantiation* of the syntactic properties exhibited by Italian, whereas English presents *no instantiation* of such properties.

The overall pattern of responses in this study is more complex and more revealing than we had anticipated. The data indicate that the English subjects have indeterminate judgments about all restructuring constructions, regardless of whether they are grammatical or ungrammatical, optional or obligatory. The judgments of the French subjects, on the other hand, are very similar to those of the Italians in the case of clitic-climbing, where there is a clear choice between the grammatical and the ungrammatical sentences, but they are radically different concerning optional change of auxiliary. The French near-native speakers seem to have acquired a divergent representation of restructuring, which allows the change of auxiliary from AVERE to ESSERE only under clitic-climbing, but not otherwise.

Interestingly, the differences between English and French informants revealed by Magnitude Estimation were not detected by Card-Sorting: once again, the latter does not contradict Magnitude Estimation but it simply obliterates the differences, producing strikingly similar patterns for the three language groups. As we saw earlier, it is plausible to interpret the divergent judgments expressed in the two experimental conditions as an effect of *time*: near-natives have metalinguistic knowledge of restructuring constructions that they can retrieve and apply in the untimed condition allowed by Card-Sorting, but not in the timed condition imposed by Magnitude Estimation.

This pattern confirms the findings of Coppieters (1987): despite the fact that both non-native groups are capable of seemingly native-like performance, their knowledge representations, particularly in the case of restructuring constructions, are substantially different from the

native knowledge representations. This suggests that similarities in performance between natives and near-natives may be a misleading indicator of near-nativeness at the competence level.

More importantly, however, this study opens new perspectives on some fundamental questions about the nature of second language acquisition.

First, undifferentiated use of the term 'near-native' conceals the fact that incompleteness and divergence are two distinct states of grammatical competence, corresponding to two qualitatively different kinds of ultimate attainment. We want to argue that the competence differences between the French and the English near-natives reflect differences in the overall representations of unaccusativity in French and English, as suggested in Chapter 3.

Second, the judgments given by the near-native subjects indicate that the use they have made of evidence in the input to modify their interlanguage grammar has been selective, despite the ample availability of positive evidence for all the grammatical aspects investigated in this study. That is, the results suggest that the availability of positive evidence of a property P in the L2 input may not be a sufficient condition for acquisition to take place. The usual assumption is that positive evidence will have the same effect on all learners, that is, it will invariably represent an instruction to the learner to adopt P, or the particular setting of P. What is not usually considered is, first, the possibility that the relevance and the salience of positive evidence may be a matter of degree and, second, that - other things being equal - some learners may be in a more favourable position to *notice* the relevance of the L2 input and to incorporate it in their developing grammar. In this study, the evidence of the relative lexical-semantic markedness of different types of unaccusative and unergative verbs was appropriately "taken in" by both French and English subjects. Evidence of optionality or obligatoriness of auxiliary change in syntactically marked restructuring constructions, on the other hand, was incorporated in a divergent representation by the French subjects and went mostly unnoticed by the English subjects.

This means that the interlanguage representations constructed by the two subject groups reflect specific features of the L1, though not in any direct way. That is, the French subjects' transformation of an optional choice into a categorical rule that generalizes AVERE is entirely consistent with the fact that restructuring constructions are peripheral cases for ESSERE-selection and therefore more permeable to the ESSERE--->AVERE substitution. But if the overgeneralization of AVERE were seen as a case of straight transfer, it would be difficult to explain why the French subjects do not adopt it for the clitic-climbing construction as well.

Similarly, it would be problematic to explain why the English subjects, unlike the French, do not overgeneralize AVERE (the only auxiliary allowed by their L1), which is what would be expected if they were adopting a simple transfer strategy: rather, they have indeterminate judgments on both auxiliaries.

The propensity of certain learners to notice, or fail to notice, the occurence of a given property P in the L2 input appears to be related to the status of the learner's native language with respect to that property. Yet this is not simply a matter of whether a property P is instantiated in the learner's native language; it is crucial to look beyond P itself, at the whole grammatical subsystem of the native language to which P belongs. Depending on the cross-linguistic characteristics of such a subsystem and on its degree of typological consistency with the L2, a given L1 may represent a more or less favourable starting point for the acquisition of P. In the case at hand, the competence differences between French and English subjects reflect not so much the presence in the L1 of the *specific* properties investigated, or the presence of these properties in the L2 input, but rather the overall representations of unaccusativity in French and English. It must therefore be recognized that both availability vs non-availability of evidence in the L2, and instantiation vs non-instantiation of a property in the L1, are continuous rather than dichotomous dimensions, and relative rather than absolute.

To conclude, we are in a position to reject the null hypothesis that there are no differences between French and English near-native speakers of Italian with respect to the acquisition of syntacticaly-based properties of auxiliary selection; French near-native speakers have divergent intuitions manifested in determinate judgments, while English near-native speakers have incomplete intuitions manifested in indeterminate judgments.

10.5 Prospect

Let us return to the starting point of this thesis:

> "In actual practice, linguistics as a discipline is characterized by attention to certain kinds of evidence that are, for the moment, readily accessible and informative: largely, the judgments of native speakers. Each such judgment is, in fact, the result of an experiment, one that is poorly designed but rich in the evidence it provides....the judgments of native speakers will always provide relevant evidence for the study of language...although one would hope that such evidence will eventually lose its uniquely privileged status."
> (Chomsky 1986: 36-37)

In this study we hope to have shown that acceptability judgments need not be elicited only in "poorly designed" experiments. Furthermore, while we agree with Chomsky that linguists would do well to make use of other sources of evidence, we have argued that when elicited in appropriately designed studies acceptability judgments can provide evidence about internal grammars that is far richer even than commonly supposed. We hope that future studies of the sort reported here will make it possible to investigate a wide range of grammatical phenomena and perhaps even to justify the "uniquely privileged status" of acceptability judgments for some time to come.

FOOTNOTES

Chapter 1

1. Gleitman and Gleitman (1979: 122), for example, argue that "...judgmental functions in language are separate from the language functions both on descriptive grounds (the data of linguistic judgments do not organize the findings for speech and comprehension in real time) and on developmental grounds (the presumed metafunctions are developmentally late to appear)."

2. Quirk and Svartvik's (1966) study (expanded by Greenbaum and Quirk 1970) experimented with "a technique for establishing degrees and kinds of acceptability in English sentences" (p. 97), consisting of two complementary tests: the Judgment Test and the Operation Test. The two tests together provide a measure of control over the basis of the informant's reaction: while the Judgment test taps preferences directly, the Operation test elicits preferences indirectly, by asking informants to perform a number of changes on sentences, such as turning the sentence into the past tense, or from positive to negative, or into a question. Quirk and Svartvik's conclusion is that the results indicate a correlation between the Operation and Judgment tests, in the sense that sentences that had a high Operation success rate tended to be accepted as normal and well-formed in the Judgment test, and sentences that had a low Operation success rate also had a low acceptance rate. However, there was much more variation in the Judgment test ("...throughout the entire range between unanimous acceptance and unanimous rejection", p. 98) than in the Operation test.

3. A well-known example of variability in judgments induced by the context of presentation is Greenbaum's (1976) study on the negation of *dare*. Three variants of the negative of *dare* were presented in sentences that were otherwise identical: *didn't dare* (without *to*), *dared not*, and *didn't dare to*:

> We didn't dare answer him back
>
> We dared not answer him back
>
> We didn't dare to answer him back

Each sentence appeared twice, paired separately with each of the other variants. While the highest ratings were given to the sentence with *didn't dare to* and the next highest to the sentence with *dared not*, judgments on the least acceptable variant fluctuated according to which of the other two variants appeared next to it: it was judged less acceptable when it was juxtaposed to *didn't dare to* than when it was juxtaposed to *dared not*.

4. The modular view of language within cognition holds that grammar and discourse context are two distinct but interacting domains (see Newmeyer 1983: 57 on this issue): according to this view, "the grammaticality of a sentence is most profitably regarded as independent of its discourse context". The problem with using contextualized sentences in experiments on linguistic acceptability is that very

little is known of the modes of interaction of grammatical knowledge and contextual variables: it would be therefore difficult to evaluate which of the two factors is at the source of judgments.

5. The rationale behind the mirror manipulation experiment was that informants who are in a state of objective self-awareness (induced by the presence of the mirror) adopt a "communicative strategy" which emphasizes semantic similarity, whereas informants who are subjectively self-aware use a "sentential strategy" that lead them to focus on the structural properties of a sentence. Carroll, Bever and Pollack's conclusion is that these strategies may be indications of "relatively stable mental schemata that underlie manifest intuitions of similarity".

6. See Cowart's study of subjacency sentences such as *Who did the Duchess sell a portrait of?* which present an apparent conflict between their theoretical ungrammaticality and their full acceptability/intelligibility to most speakers.

7. See Pateman (1985: 50): "..insofar an adaptive rule is developed self-consciously on the basis of beliefs about desirable and undesirable linguistic forms, rather than from unconscious processing of speech input...it is plausible to suppose, from the nativist standpoint, that the process will engage the *subject* making use of cognitively resources not specific to language, either exclusively or in combination with the language faculty."

8. According to Nunnally, the word *judgment* refers to those types of responses where there is a *correct* response. The word *sentiment* , on the other hand, refers to responses concerning attitudes, personal reactions, likes and dislikes, etc, where there is not external standard of accuracy that makes sense.

9. The three working principles that, according to Labov (1975: 31), "offer a fairly sound basis for continued exploration of grammatical judgments" are clearly based on an interest in sociolinguistic variation, rather than in individual grammars:

I. The Consensus Principle: is there is no reason to think otherwise, assume that the judgments of any native speaker are characteristic of all speakers of the language.

II. The Experimenter Principle: if there is any disagreement on introspective judgments, the judgments of those who are familiar with the theoretical issues may not be counted as evidence.

III. The Clear Case Principle: disputed judgments should be shown to include at least one consistent pattern in the speech community or be abandoned.

10. An example of the principle of "letting the grammar decide" at work is the sentence *John's likelihood to win the prize*, quoted by Newmeyer (1983). This sentence is ungrammatical according to the lexicalist model developed in Chomsky (1970) and grammatical according to the transformationalist model defended by Newmeyer (1971). Newmeyer acknowledges that both he and Chomsky presented their analyses as if there were no uncertainty about the data, although the sentence in question is indeed an unclear case. But the point that Newmeyer wants to emphasize is

that the apparent disagreement derived from differences in the *analysis* of the sentence, not in differences in *intuitions*.

11. Mohan (1977) raises the interesting issue of the predicting power of Lakoff's fuzzy grammar model, i.e. the fact that it can only make the weak claim that members of a speech community will share the same ordinal scale for an acceptability hierarchy, but not the stronger claim that they will agree on the ordering of the differences between items on the hierarchy. This issue will form the central part of Chapter 2 of this thesis.

12. In Ross's (rather humorous) characterization of relative acceptability, sentences of a language are viewed by speakers as falling into three groups: a core, a bog, and a fringe.

13. Fodor (1989: 134) points out that the discontinuity assumption would make it difficult to explain how a learner copes with sentences like *John believes Bill to have left*, which are presumably not generated by a known choice of parameter values: "...once out in the periphery, unconstrained by core principles, there are a vast number of moves he might make. He might decide, for example, that a lexical NP needn't have case, or that an NP can have case assigned by a non-governor or by nothing at all...in fact all learners assume instead that some (but not all) verbs can trigger S' deletion (or perhaps, as has also been proposed, that for these verbs S' is transparent to government)". The point here is that peripheral constructions are clearly related to the core principles and this lends support to the continuity assumption.

14. Lightfoot (1991) well represents the generativist position on language change: "..gradualness exists: triggering experiences may change gradually, and new parameter setting may gradually permeate a speech community. Much of this gradual change may have no immediate effect on the setting of *structural* parameters, which are reset only when the triggering experience has changed in some critical fashion" (p. 163). As will be seen in Chapters 8-10, the evidence presented in the present study is in principle consistent with this view.

15. Critics of the generativist position on the nature of language change, such as Kroch (1989), seem nevertheless to agree with it when they claim that "change is more often gradual than abrupt...and one generation is more likely to differ from its predecessor in the frequency with which its speakers use certain forms than in whether those forms are possible at all. *Only when the frequency of a form drops below a minimum threshold do learners reanalyze their grammatical systems so as to exclude it* (our emphasis)".

16. Interlanguage indeterminacy due to ignorance was for a long time the only kind of indeterminacy recognized by second language acquisition researchers. Typical of this early attitude is the study by Schachter, Tyson and Diffley (1976).

17. Much of second language acquisition research conducted in the past five years has been concerned with the 'availability of Universal Grammar' issue. The aim of such research is to achieve an explanation of how adult learners attain competence in a second language, and current linguistic theory offers a natural framework for this purpose. Arguments in favour of a role for UG in L2

acquisition centre around the logical problem of L2 acquisition, i.e. the fact that learners who achieve a reasonable degree of success in the second language have necessarily 'gone beyond' the input they were exposed to. Open questions are whether UG principles are available 'in toto', or whether they are available only through the specific instantiations exhibited by the L1; whether interlanguage grammars are 'natural grammars' (i.e. constrained by UG) at every stage of development; what kind of input contributes to L2 developmental stages. A substantial number of researchers, on the other hand, do not believe in the availability of UG, and emphasize the fundamental differences between L1 and L2 acquisition, the difficulties faced by L2 learners, and the impossibility of 'complete' acquisition by adult learners (see the section on 'near-nativeness' below).

18. Early work on second language learnability focused on the availability of the Subset Principle to L2 learners (see White 1989, 1992 for a comprehensive review). The Subset Principle is a learning principle (available to the L1 acquirer) which guarantees acquisition without negative evidence: the principle is an in-built instruction to the learner to be conservative, preventing her from adopting over-inclusive parameter-settings in the absence of explicit evidence. The findings on L2 acquisition indicate that the principle is not available to L2 learners: this, together with the natural tendency of learners to transfer the parameter settings instantiated by their L1s, explains the learning asymmetries that have been observed between learners whose L1 is a superset of the L2 and learners whose L1 is a subset of the L2: only in the former case does the learner have a potential problem of overgeneralization. Recent developments in linguistic theory, however, suggest that parametric variation may be the result of variation in the properties of lexical items rather than in the principles of Universal Grammar (Wexler and Manzini 1987; Ouhalla 1991). A further refinement of this 'lexical parameterization hypothesis' identifies the site of variation in the properties of the functional categories rather than in the major lexical categories ('functional parameterization hypothesis'). The implication for second language acquisition is that parameter-resetting may be an impossibility for the L2 learner because functional categories may be acquired by the child as a result of a genetically determined schedule (Tsimpli and Roussou 1991).

19. Indeterminacy induced by input restrictions has even more obvious consequences in situations of language attrition, where it is the L1 which is removed from its natural context. The absence of L1-input, coupled with the overwhelming presence of L2-input, leads to a progressive 'disintegration' of the core through a process of constant erosion which undermines the determinacy of its structures. This is what Py (1986) defines as the process of 'positive retroaction' ('rétroaction positive') which accompanies the growing indeterminacy of norms, due to the increasing permeability of the core to L2-inspired deviations. Because of the reduced availability of UG constraints, there is seldom simple replacement of L1-core features with L2-core feature: more often there is no replacement at all, with resulting indeterminacy, variation, and increased openness to change.

20. The critical test for the plausibility of a 'logical problem' of second language acquisition is that L2 learners should be able to acquire complex syntactic properties that (a) are not instantiated in the L1 and (b) that cannot be the focus of explicit classroom instruction (White 1989).

21. Plato's problem: how come we know so much from so little evidence? Orwell's problem: how come we know so little in spite of so much evidence? (Chomsky, 1986; Hale, 1988).

22. Birdsong (1991) in a study that partly replicated Coppieters' but also tested other linguistic variables, found significant difference between natives and near-natives, using cumulative deviance from native norms: however, he also found that a number of individual near-natives performed well within the native range. It is important to see these findings, and the results of the present study, as complementary rather than incompatible: ultimate attainment may be very different depending on the aspect of grammar acquired, and learners may be able, in certain circumstances, to reach native-like competence with regard to some aspects, but not others.

23. Cross-sectional designs have the advantage of being convenient to implement and of allowing the simultaneous testing of a sizable sample of subjects. However, one has to bear in mind that they are only a *simulation* of real-time acquisition because the data are not collected from the same individuals.

Chapter 2

1. Category scales are said to be 'unique up to a one-to-one transformation', since the classification on these scales may be equally well represented by any set of labels, and labels designating the various sub-categories may be interchanged if this is done consistently. Ordinal scales are said to be 'unique up to a monotonic transformation', since these scales allow any operation that does not alter the relative ordering of items. Interval scales are 'unique up to a linear transformation' because any operation on them must preserve not only the ordering of items but also the relative differences between items. Finally, ratio scales are said to be 'unique up to multiplication by a positive constant', which means that the ratios between any two numbers are preserved when the values on the scale are all multiplied by a positive constant. (Siegel and Castellan 1988)

2. However, not all experimentalists hold the same strict attitude against the use of parametric statistics with data that do not fully satisfy the parametric assumptions (see for example Nunnally 1967).

3. Categorical responses may have a role in studies of sentiments, where it is important to learn the absolute level of responses to stimuli (for example, in studies of attitudes towards different national groups, for which it has been hypothesized that people near the 'neutral' point are more susceptible to change than people who are distant from the neutral point in either direction).

4. The present author has shown (Sorace 1985a) that the findings of a number of studies can be read in different ways according to how inconsistency is interpreted. In White (1985), for example, most inconsistent responses in a binary judgment test were produced by intermediate learners: if these learners had not been constrained by a dichotomous rating scale, they may have been able to express uncertainty, and it would have been possible to determine the direction in which their interlanguage competence was changing.

5. This tendency could be defined the 'may-be-saying factor' in analogy with the well-known 'yea-saying factor' in acceptability judgments, i.e. of a tendency to accept sentences regardless of their particular structure (Mohan 1977).

6. Multiple-point scales can yield even interval information if subjects are instructed to use the scale as if the distances between successive categories were the same: this can be obtained, for example, by labelling categories in terms of percentages. The problem with this method of *equal-appearing intervals* is that the experimenter imposes her own categorization on the subjects, without any guarantee that the subject's categorization would coincide with her own.

7. Kellerman's study was concerned with the inter-relationship of the different meanings of *break* and *breken* in the intuitions of native speakers of English and native speakers of Dutch. By asking subjects to sort nouns typed separately on cards into piles according to *similarity of meaning*, Kellerman was able to determine a 'semantic space' where judgments of inter-relatedness are made, and the dimensions that may be responsible for variation in responses.

8. The use of logarithms allows multiplication, division and exponentiating to be replaced by the simpler operations of addition, subtraction and multiplication (Lodge 1981).

9. The general formula for the geometric mean is:

$$\text{Geometric mean} = \sqrt[n]{x_1, x_2, x_3 \ldots . x_n}$$

10. According to Poulton (1986, 1989), the possible biases affecting magnitude estimation can be grouped under the following categories:

(a) biases produced by telling the informants what to do (including initial demonstrations, examples, instructions to judge ratios);

(b) biases induced by the experimenter's choice of the range of stimuli to present, the spacing of the stimuli, and their relative frequencies;

(c) biases produced by the absence or presence of a standard stimulus paired with each variable stimulus, and the position of the stardard in the range of stimuli.

(d) biases induced by decisions about the choice of the direction and distance of the very first variable from the standard.

Chapter 3

1. Extensive treatments of the foundations of Relational Grammar can be found in Perlmutter (1983) and Perlmutter and Rosen (1984). For interesting work on French unaccusativity in a Relational Grammar framework, see Legendre (1989a, 1989b, 1990). For a behind-the-scene history of the Unaccusative Hypothesis from its origins in Relational Grammar to its more successful version in Government and Binding, see Pullum's essay "Citation etiquette beyond Thunderdome" in Pullum (1991).

2. We adopt in the present study the 'standard' version of Government and Binding theory presented in Chomsky (1981), which is also at the basis of Burzio's work. We do not take into account more recent developments concerned with the 'split-INFL' hypothesis (Pollock 1989), with the hypothesis that subjects are generated under the VP node (Sportiche 1988), or with the 'functional parameterization hypothesis' (Ouhalla 1991), all of which may in the future inspire new insights on the syntactic characterization of auxiliaries. An alternative accounts of auxiliary selection within generative grammar is provided by Guéron and Hoekstra (1988).

3. In fact the possibility of non-assignment of subject theta-role is exhibited by various (but not all) verbs that take sentential complements, such as Raising verbs, but it cannot be related in any way to the presence of the sentential complement, or with the subcategorization frame. This parameter may therefore be expected to vary among verbs subcategorized for sentential complements, and the variation to be governed by lexical, not by syntactic factors, in the same way as for verbs subcategorized for NP objects.

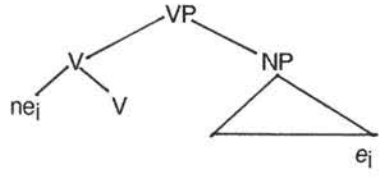
4. Burzio builds evidence for the generalization in two steps (see Chapter 3 of his book). First, he establishes that only verbs that can assign theta-role to the subject can assign accusative case to an object (which can be expressed as $-\theta_S$---> -A, where 'θ_S' refers to the property of a verb to assign theta-role, and A refers to the property of the same verb to assign accusative case). This is relatively easy to demonstrate empirically: for example, NPs to the right of unaccusative verbs cannot appear in the accusative case (*Sarei soffocato anche *me / io* 'would have choked also me/I') unlike NPs to the right of their transitive alternants (*Il caldo avrebbe soffocato anche me / *io* 'The heat would have choked also me/I'). Then, Burzio shows that the other side of the generalization, that is, that only the verbs that can assign accusative case to their object also assign theta-role to the subject position (which can be represented as -A --> $-\theta_S$.). This time, the statement is true on theoretical, rather than empirical grounds: if a verb that takes a direct object fails to assign case to it, then it will <u>have</u> to fail to assign theta-role to the subject position because the direct object will have to either (a) be linked to a non-argument subject, or (b) move into subject position. Both possibilities require that the subject is not assigned theta-role. Burzio therefore concludes that the stronger statement derived from a conjunction of the two conditions is actually

true, namely that θ <---> A: <u>all and only</u> the verbs that can assign theta-role to the subject can assign accusative case to an object.

5. We will not consider here the third piece of evidence employed by Burzio in support of his analysis of unaccusativity: the different linear order of constituents with unaccusative and unergative verbs ("inversion"). Essentially, the i-subject of an unaccusative verbs tends to precede certain complements of the verb, while the i-subject of an unergative verb tends to follow them (as in *Interverrà Paolo a risolvere il problema* 'Paolo will intervene to solve the problem' vs *??Sperava Paolo di risolvere il problema* 'Paolo was hoping to solve the problem', where *intervenire* is unaccusative and *sperare* is unergative).

6. However, Burzio's arguments against the movement analysis are not very strong, as he himself recognizes (see p. 33-36).

7. The structure of (40b) would be



where the clitic c-commands its trace.

8. The classic definition of government (Chomsky 1981) can be expressed as follows:

α governs γ in a configuration like

[β .... γ .... α .... γ ....]

where: (i) α = X^0 (= a lexical element)

(ii) where φ is a maximal projection, if φ dominates γ then φ dominates α

(iii) α c-commands γ.

9. Belletti and Rizzi propose an amendment to the above definition which has the effect of allowing the government relation to hold between V and all the lower projections of NP.

This amendment affects clause (ii) of the above statement:

(ii) where φ is a maximal projection, if φ dominates γ, then *either* φ dominates α, *or* φ *is the maximal projection of* γ.

This revised definition allows PRO to be governed by V in (35b) because the boundaries of the object NP are transparent to government of the head, and therefore rules out this sentence.

10. Notice that sentences such as (37b) (*Sono passati due*) become acceptable if they are pronounced with an intonational break between the verb and the post-verbal subject. Belletti and Rizzi argue that the effect of an intonational break has a structural correlate: the presence of a break indicates adjunction of the subject to S, whereas the absence of a break indicates the attachment of the subject to VP.

11. This process can be represented as follows:

In a structure like

[...dummy$_i$....NP$_i$....]

1 2

where 1 c-commands 2:

a. copy Case of 1 on 2

b. transmit theta-role of 1 to 2.

12. The fact that dummy elements cannot keep a thematic relation is shown by the ungrammaticality of sentences like **there came*.

13. Note the conceptual similarity between the binding relation posited in Government and Binding to account for ESSERE-assignment and the mechanism of Multiattachment that, as we saw earlier, accounts for the phenomenon in Relational Grammar.

14. Analogous restructuring phenomena for Spanish are the object of a study by Aissen and Perlmutter (1983) which, although framed in Relational Grammar, refers to Rizzi's position.

15. The schematic representation of the Specified Subject Condition is

....X....[$_S$...Y...].....X.....

where no rule can involve X and Y if S has a subject distinct from Y and not controlled by X.
As Rizzi observes (p. 40), in the examples in (60) Y is the base position of the pronoun, and X is the clitic position.

16. Rizzi notices (and rejects) the counterargument that all these syntactic phenomena are blocked not by structural differences determined by the application of restructuring, but by simple ordering conditions on rules. We shall not examine the details of his analysis here (see Rizzi 1982: 8-9).

17. In fact the asymmetry in the relation between the two phenomena points to the ill-understood nature of clitic-climbing in general. As Burzio notices, the ungrammaticality of a sentence like **Questi libri si vorrebbero proprio dargli* 'These books, one would really want to give them to him' (where clitic-climbing has not applied) contrasts with the grammaticality of sentences like *Paolo sarebbe voluto andarci* 'Paolo would have liked to go there', where restructuring does not trigger clitic-climbing. In other words, restructuring seems to be a necessary, but not a necessary and sufficient condition for clitic-climbing to occur.

18. This point can be appreciated when longer sequences of restructuring verbs are considered, like *Maria li avrebbe potuti stare per andare a prendere* 'Maria would have been able to be on the point of going to get them herself', where the last verb of the sequence selects AVERE and there is no change of auxiliary, compared to *Maria ci sarebbe dovuta cominciare ad andare* 'Maria would have had to start going there', where the last verb selects ESSERE, and thus restructuring takes place. Rizzi's rule is descriptively adequate, but it does not provide an explanation of the directionality of change.

19. Interestingly, Burzio's main arguments are concerned with the selectional restrictions observed by verbs in restructured and non-restructured constructions: essentially, the fact that the semantic

interpretation of a construction is not affected by the application of restructuring, which runs against what any base-generation analysis would predict.

20. Again, the arguments brought forward by Burzio in defense of the non-deletion of the embedded subject under restructuring have a semantic flavour. The interpretation of a non-restructured sentence like *Maria vuole prenderlo* "Maria wants to take it' is the same as the interpretation of its restructured counterpart *Maria lo vuole prendere* : in both cases, *Maria* is understood as the subject of both main and embedded verbs (a fact that is captured by the Control analysis of these constructions). Deletion of the embedded subject would have the undesirable effect of eliminating information crucial to the semantic interpretation.

21. See footnote 17 above.

22. As Hoekstra (1984) points out, this asymmetry confirms that ESSERE is the marked member of the auxiliary pair.

23. "It would seem that what is parameterized is the notion of government that enters into the system. At least two different notions seem to be needed to appropriately distinguish the governed NP in (b) from the one in (c), both internally to each language, and across languages. We may refer to them as STRONG GOVERNMENT, obtaining in (b), and WEAK GOVERNMENT, obtaining in (c). Perhaps the relation between the verb and the clitic in (a) can also be captured in terms of government: a third and the strongest notion (call it SUPERGOVERNMENT)". (Burzio 1986: 140)

24. Legendre (1989) lists nine diagnostics of unaccusativity: object raising, *croire* constructions, participial equi and absolute, reduced relatives, cliticization of the embedded indirect object in causative *faire* constructions, parallel transitive structures, auxiliary selection, nominalizations, and stativity.

25. Two of the examples in (89) have a controversial acceptability status. (89b) is obviously grammatical in the irrelevant reading resulting from 'il' being personal, rather than impersonal (i.e. 'he only sank three of them'). (89d) is grammatical in the agentive meaning 'the boat that was sunk yesterday', which is incompatible with the insertion of the adverb "all by itself". These ambiguities may be responsible for the fact that (89b) and (89d) elicit disagreement among French native speakers. Such lack of consistency confirms the inherent indeterminacy of paired unaccusatives with a transitive alternant in French.

26. Labelle in fact argues that there are both syntactic and semantic differences between paired unaccusatives in reflexive form and paired unaccusatives in non-reflexive form (she uses the terms 'reflexive' and 'intransitive'). She assumes "a strict correspondence between the argument structure of lexical items and their conceptual structure" (p. 307). The main semantic distinction she draws is that between changes of state taking place autonomously, and changes of state brought about by some external factor. The subject of a reflexive verb is not actively responsible for the unfolding of the event, whereas the subject of a non-reflexive verb is "both the motor and the locus of the change". It follows that non-reflexive verbs have

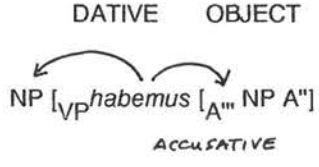
an argument that qualifies as a thematic subject but reflexive verbs have an argument that qualifies as a thematic object. Notice that this semantic characterization is similar to that usually applied to the unergative/unaccusative distinction (see section 3.8 in this chapter). Labelle's conclusion in fact assimilates non-reflexive change-of-state verbs to the class of unergatives.

27. See Burzio's (1986: 163): "...*there* is *semantically* constrained to presentational contexts, or verbs of appearance" and Haegeman (1991:310): "...unlike verbs of movement and (change of) state, (verbs such as) *sink* do not appear in the *there* construction.". Cf also Napoli (1988: 136), who claims that both unaccusative and unergative verbs may appear in *there* sentences, provided that they "function primarily to establish the existence or presentation of its GF subject in the discourse or in some location (spatial or temporal)". Notice that Haegeman, like Labelle for French, excludes verbs with a transitive alternant from the class of unaccusatives.

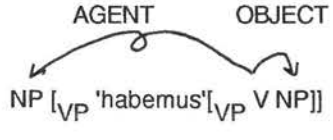
28. The semantic generalizations underlying verbs of motion are correlated to syntactic properties such as the ability to assign accusative Case. The *run* -type verbs differ from the *arrive*-type and the *roll*-type in that they can be followed by bare NPs, confirming their ability to assign case (thus, *They swam the English Channel* is possible but **The ball rolled the lane* and **they arrived the airport* are not. Furthermore, the *run*-type can appear with expletive objects (*He ran the shit out of his shoes*) but the *roll*-type cannot (**The wagon rolled the shit out of its wheels)*

29. Analyses of unaccusativity in non-Romance languages from a GB perspective have been written for German (Grewendorf 1988) and Dutch (Hoekstra 1984).

30. Salvi (1977) suggests the following structural analysis for the classic Latin construction in (107a):



where the verb *habemus* governs a Small Clause with an adjectival head. The verb assigns a thematic role to its subject (DATIVE) and accusative Case to its object, and to the subject of the Small Clause. In contrast, (107b) has the following structure:



where the verb has lost the possibility of assigning accusative case to the direct object, the subject of the small clause cannot be filled, and the small clause is therefore eliminated.

31. Vincent (1982) avails himself of the framework of Case Grammar (Fillmore 1968), at the basis of which there is a highly rigid idea of the correspondence between syntactic and semantic roles: every semantic role (Agent, Experiencer, Patient, Theme etc) is claimed to map into a unique 'deep syntactic case'. As Jackendoff (1991) points out, Case Grammar is one of the earliest statements about *linking rules* (which, as we will see later in this chapter, occupy a central role in recent theoretical debate about argument structure. One of the main problems in assuming a tightly constrained theory of the mapping between syntax and semantics is that surface grammatical relations do not obey such a rigid correspondence: therefore movement rules, or other kinds of syntactic transformations are posited to account for surface syntactic distribution. Less restrictive linking theories have been proposed that do not assign each thematic role to a particular syntactic position: instead, they assume a hierarchical mapping between an ordered list of thematic roles and an ordered list of syntactic relations. We will examine a number of these hierarchical approaches in 3.8 below.

32. Antinucci and Miller's data have recently been reanalysed in a principles and parameters framework by Borer and Wexler (1992). This analysis is based on the idea that Italian children who produce past participle agreement with direct objects analyse participle phrases with direct objects as adjectival phrases (APs) rather than as transitive VPs, and overegeneralize Spec-Head agreement to the relation between the participle and the object (notice the convergence of part of this analysis with Salvi's treatment - mentioned earlier in footnote 30 - of the Classical Latin construction with *habere*). The child's error derives from the existence of a maturationally determined Unique External Argument Proto Principle (UEAPP), which requires that every predicate be associated with a unique external argument. This principle, according to Borer and Wexler, account for the almost complete absence in children's speech of unergative verbs in the present perfect: the presence of two predicates (the auxiliary AVERE and the past participle) would violate the UEAPP in child grammar. The principle also provides an alternative account for the fact that the initial setting of the null subject parameter allows null subjects in non-null-subject languages like English (details will be omitted here: see Borer and Wexler 1992: 183).

33. The explanation for NE-cliticization in Role and Reference Grammar exemplifies the hierarchical approach to linking theory mentioned in footnote 31. The reason why NE-cliticization correlates with transitive verbs is that the direct object of a transitive verb is the lowest-ranking argument in the state predicate in the logical structure of the verb in terms of the Actor/Undergoer accessibility hierarchy in Fig. 3.6; it correlates with unaccusative verbs, where the lowest-ranking argument in the state predicate in the verbs's LS is the subject. NE-cliticization is not allowed by ESSERE itself (although ESSERE selects ESSERE as an auxiliary): in a sentence like *Molti libri sono interessanti* the theme argument is the attribute and the locative argument is the bearer of the attribute. As the theme is incorporated into the

predicate and therefore not available as a syntactic argument in the clause, a marked linking takes place that realizes the locative as the undergoer subject. If NE-cliticization is attempted on the clause (**Ne sono interessanti molti*), the result is ungrammatical because NE does not realize the lowest-ranking argument in the state predicate.

34. Elsewhere it is stated, however, that ESSERE is the unmarked auxiliary in distributional terms (whereas in French ETRE is marked) because if one considers all the major types of constructions in Italian ith respect to the parameter of transitivity and basic vs derived status, the following pattern emerges:

AVERE: underived transitive constructions
intransitive activity verbs

ESSERE: all others

35. The various theories differ in terms of they specify the syntactic level of representation. As we have just seen, Role and Reference Grammar maps directly from the semantic hierarchy onto phrase structure, and so do the early approaches to the linking problem, such as Fillmore's Case Grammar; as we will see,Levin and Rappaport (1988), Zubizarreta (1987), and Pinker (1990) map onto one (or more) levels of argument structure that interface with syntactic configurations. Jackendoff's most recent approach presents a formal theory of conceptual semantics where lexical entries carry all the syntactically relevant information as annotations at the level of conceptual structure: that eliminates the need for a separate level of argument structure (see Jackendoff 1991, chapter 11).

36. Depending on the specific theory of argument structure, the term 'linking rules' refers either to the mapping from lexical-semantic structure onto argument structure or to the mapping from argument structure onto syntactic structure.

37. Levin and Rappaport's linking rules are an example of the hierarchical approach to the syntax/semantic interface. The rules map the ordered thematic roles in lexical-conceptual structure onto variables in argument structure from left to right.

38. As Roberts (1990) points out in his review of Zubizarreta's book, this approach is closer "both in letter and spirit" to GB theory than any of the other theories of lexical representation. However, the lexical-syntactic level does not have an obvious counterpart in GB and, according to Roberts, may not be entirely justified from a theoretical point of view since other modules of the GB model do most of the work that Zubizarreta attributes to the L-R level.

39. Zubizarreta makes a rather vague suggestion that "the notion is perhaps a remnant of a primary semantic category in early stages of language acquisition. In the adult grammar it probably has the status of a 'recessive' feature to which rules and principles of grammar do not refer, but which proves useful in stating certain selectional restrictions."(p. 12).

40. L-R mediates only the mapping of verbs and adjectives (defined "direct arguments" from the semantic level onto syntax. Nouns ("indirect arguments") bypass L-R entirely because

they cannot be L-predicates (and therefore they cannot have a lexical representation) for semantic reasons: they can be R-predicates (those which are predicated of a reference). The peculiar status of nouns explains the differences between nominals and their clausal counterparts (details of which will be omitted here: see chapter 2 in Zubizarreta's book).

41. The Default Linking Rule creates an NP-structure of the following type:

$Mary_i$ $[_{VP}{}^{y}$ arrived^{y} $e_y]$

The Rule of Predication (whose main function is to assign a value to the variable index borne by the predicate) equates *i* and *y* in the above structure. These two rules do the work of NP-movement, because they externalize the internal argument and establish a binding relation betwen the surface subject and the trace in object position. Within this system, NP/trace relations are the projection onto syntax of *lexical binding relations* (Zubizarreta 1987: 16).

42. However, the relation between transitive and unaccusative verbs is reversed in the rule of *causativization*, which derives a causative verb from an adjective (i.e.*modern* --->*modernize*). In order to collapse the rule of anti-causativization and the rule of causativization, Zubizarreta assumes that anti-causatives are the basic lexical items from which causatives are derived, rather than vice versa. Thus, transitive *sink* would be derived from unaccusative *sink* via insertion of the CAUSE feature: the insertion of CAUSE would triggers the internalization of the original external argument and the insertion of a new external argument. This view, however, is not without problems. As Roberts notes, it is not clear why the insertion of CAUSE does not lead to a sentence like *The ship sank* meaning 'the ship caused something to sink'. Zubizarreta strives to avoid reference to thematic roles, but it would seem that in order to avoid this result one would have to specify that the insertion of CAUSE entails the insertion of a 'CAUSER': this move would be empirically equivalent to making direct reference to a thematic role 'causer' (for this and other problematic aspects of Zubizarreta's use of the CAUSE and STATE semantic notions, see Roberts (1990: 484).

43. The standard definition of *external argument* relies on the idea that the external argument is realized outside the maximal projection of the predicate, the d-structure subject for a verb. As Grimshaw points out, what is not explained is why there should be an argument distinguished in this way, and why there is only one such argument. Grimshaw's proposal makes the notion of external argument a relational notion, dependent on the relationships among other arguments. Her proposal has the further advantage of bringing the notion of external argument in line with recent developments in linguistic theory, which argue that subjects are generated under the VP node (and then move to Spec of IP) rather than being generated under IP (Grimshaw does not elaborate this point).

44. Cf. forms such as *sò bingút* in many Catalan dialects.

45. Some examples (from Gougenheim 1951) are:

C'estoit vrayment un Ange Qui pour nous prendre estoit vollé des cieulx (Ronsard, *Amours*, 1552)

C'estoit la cause pour laquelle ils estoyent courus (Amyot, Démosthène, 26).

Je n'y fusse pas esté jusques à minuyt (Nicolas de Troyes, *Grand Parangon*, 18)

46. There seem to be some constraints, however, on auxiliary selection with weather verbs. We saw in 3.8.3.5 that when they are turned into telic verbs (*Mi è piovuto sulla testa*) they admit only ESSERE. The thematic features of the subject also appear to lead to preferences for one auxiliary over the other. *Decollare* 'take off', for example, selects either auxiliary if the subject is non-agentive (*L'aereo é / ha decollato* 'The plane has taken off') is more acceptable with AVERE if the subject is agentive (*Il comandante ha / ?è decollato* 'The captain has taken off').

47. The only exception in this category is the verb *rester*, which takes ETRE.

48. This hypothesis attributes a central role to linking rules, along the lines indicated by Jackendoff (1991). Linking rules are necessary because they articulate the correspondences between the syntactic and the semantic components, each of which functions autonomously on the basis of its own principles. Once one acknowledges the necessity of the linking component, it is only a short step to assume the 'autonomy of linking rules', i.e. the idea that linking rules have their own properties and typology defined in conjuction with, but not necessarily dependent on, the syntactic structures and the conceptual structures they relate. This view is compatible with modular theories of cognition (see the Epilogue in Jackendoff 1991), and opens up new perspectives on theories of learnability (see Pinker 1991 for the acquisition of linking rules and argument structure in first language acquisition; virtually no work has been done on the acquisition of argument structure in a second language).

Chapter 5

1. As was pointed out earlier, recent evidence from second language acquisition research suggests that the Subset Principle may not be available to adult learners.

2. In this pilot study, as well as in the main study, not all the verb categories on the Unaccusative Hierarchy in Table 3.8 were included in the research design: in Pilot 1, the two categories of 'change-of-location' and 'change-of-condition' were collapsed into a single category 'change-of-state'. In the main study, the category 'change-of-condition' was not included in the design. It was felt that, in this first attempt to investigate the reality of these unaccusative sub-types in the intuitions of native and non-native speakers, it was more crucial to establish the distinction between 'change' and 'absence of change' than that between change of location and change of condition.

3. Striking evidence for the 'naturalness' of the connection between essere and unaccusativity comes from Zobl (1989), who reports that learners of English from different language backgrounds (including non-Romance) produce errors involving the use of be as the auxiliary of unaccusative verbs (i.e. 'the most memorable experience in my life was

4. In order to determine the general sensitivity of the three methods to degrees of acceptability, this analysis compared verb types across the three broad classes of unergative, unaccusative and transitive.

5. The analysis of perfectivity and imperfectivity we present here relies on Mereu (1981).

Chapter 6

1. We follow the procedure established in Clark (1973). Min F' values provide a more robust test on significance that complements the ANOVA: it combines both F_1 values (significant effects in a by subject ANOVA) and F_2 values (significant effects in a by materials ANOVA) and uses the following formula:

$$\text{Min } F'_{(i,\,j)} = (F_1 F_2) / (F_1 + F_2)$$

where: F_1 = the by subject F for a particular effect

F_2 = the by-materials F for the same effect

i, j = numerator and denominator degrees of freedom for determining the significance of Min F' on standard F tables

i = the numerator degrees of freedom for the effect

$$j = \frac{(F_1 + F_2)^2}{F_1^2 / n_2 + F_2^2 / n_1}$$

n_1 = the denominator degrees of freedom on F_1

n_2 = the denominator degrees of freedom on F_2

Chapter 10

1. A research project on "The validation of magnitude estimation of linguistic acceptability" funded by ESRC is currently under way at the University of Edinburgh. The present author is one of the three principal investigators.

BIBLIOGRAPHY

Adjémian, C.1976. On the nature of interlanguage systems. *Language Learning* 26: 297-320.

Adjémian, C. 1982. La specificité de l'interlangue et l'idealisation des langues secondes. In Centre de Recherche, Université de Paris VIII (ed.), *Grammaire transformationelle: théorie et méthodologie.* pp. 421-439.

Aissen, J.L and Perlmutter, D.M. 1983. Clause reduction in Spanish. In D.M. Perlmutter (ed.), *Studies in Relational Grammar 1*. Chicago: The University of Chicago Press.

Altmann, G.T.M and Steedman, M. 1988. Interaction with context during human sentence processing. *Cognition* 30: 191-238.

Anderson, S. 1971. On the role of deep structure in semantic interpretation. *Foundations of Language* 6: 387-396.

Anderson, S.R. 1977. Comments on the paper by Wasow. In P. Culicover, T. Wasow, and A. Akmajian (eds.), *Formal Syntax.* New York: Academic Press.

Antinucci, F., and Miller, R. 1976. How children talk about what happened. *Journal of Child Language* 3: 167-189.

Arditty, J. and Perdue, C. 1982. Variabilité et connaissance en langue étrangère. *Encrages: numero special de linguistique appliquée* : 32-43.

Baker, M. 1983. Objects, themes, and lexical rules in Italian. In L. Levin, M. Rappaport and A. Zaenen (eds.), *Papers in Lexical-Functional Grammar.* Bloomington: Indiana University Club.

Belletti, A. and Rizzi, L. 1981. The syntax of *ne:* some theoretical implications. *Linguistic Review* 1.

Benveniste, E. 1968. *Problèmes de linguistique générale*. Paris: Gallimard.

Benzing, J. 1931. Zur Geschichte von *ser* als Hilfszeitwort bei den intransitiven Verben im Spanischen. *Zeitschrift für romanische Philologie* 51: 385-460.

Berruto, G. 1987. *Sociolinguistica dell"Italia contemporanea*. Firenze: La Scuola.

Birdsong, D. 1991. Ultimate attainment in second language acquisition. Paper presented at the 16th Boston University Conference on Language Development.

Birnbaum, M.H. 1980. Comparison of two theories of "ratio" and "difference" judgments. *Journal of Experimental Psychology: General.* 109: 304-319.

Borer, H. and Wexler, K. 1992. Bi-unique relations and the maturation of grammatical principles. *Natural Language and Linguistic Theory* 10: 147-190.

Botha, R.P. 1973. *The Justification of Linguistic Hypotheses.* The Hague: Mouton.

Bresnan, J. 1982. *The Mental Representation of Grammatical Relations.* Cambridge MA: MIT Press.

Burzio, L. 1986. *Italian Syntax. A Government-Binding Approach.* Dordrecht: Reidel

Canale, M., Mougeon, R. and Bélanger, M. 1978. Analogical leveling of the auxiliary être in Ontarian French. In M. Suner (ed.) *Contenporary Studies in Romance Linguistics.* Washington: Georgetown University Press.

Cazden, C.H., Cancino, H., Rosansky, E.and J. Schumann. 1975. *Second Language Acquisition Sequences in Children, Adolescents and Adults.* United States Department of Health, Education and Welfare.

Centineo, G. 1986. A lexical theory of auxiliary selection in Italian. *Davis Working Papers in Linguistics*. 1: 1-35.

Chaudron, C. 1983. Research on metalinguistic judgments: a review of theory, methods and results. *Language Learning* 33: 343-377.

Chomsky, N. 1957. *Syntactic Structures.* The Hauge: Mouton.

Chomsky, N. 1961. Some methodological remarks on generative grammar. *Word* 17: 219-239.

Chomsky, N. 1970. Remarks on Nominalization. In R. Jacobs and P. Rosenbaum (eds.) *Readings in English Transformational Grammar.* Waltham MA: Ginn.

Chomsky, N. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, N. 1986. *Knowledge of Language: its Nature, Origins, and Use.* New York: Praeger.

Clark, H. 1973. The language-as-sixth-effect fallacy: a critique of language statistics in psycholinguistic research. *Journal of Verbal Learning and Verbal Behavior* 12: 335-359.

Coppieters, R. 1987. Competence differences between native and fluent nonnative speakers. *Language* 63: 544-573.

Corder, S.P. 1974. The significance of learners' errors. In J. Richards (ed.), *Error Analysis*: 19-27. London: Longman.

Corder, S.P. 1981. *Error Analysis and Interlanguage*. Oxford: Oxford University Press.

Cowart, W. 1989. Illicit acceptability in *picture* NPs. *Papers from the 25th Annual Regional Meeting, Chicago Linguistics Society,* pp.27-40.

Dowty, D. 1979. *Word meaning and Montague Grammar.* Dordrecht: Reidel.

Dulay, H. and Burt, M. 1973. Should we teach children syntax? *Language Learning* 23:245-258.

Eubank, L. (ed.) 1991. *Point Counterpoint. Universal Grammar in the Second Language*. Amsterdam: John Benjamins.

Felix, S.W. 1985. More evidence on competing cognitive systems. *Second Language Research* 1: 47-72.

Flynn, S. 1987. *A Parameter-Setting Model of L2 Acquisition.* Dordrecht: Reidel.

Flynn, S. and O'Neill, W. 1988. *Linguistic Theory and Second Language Acquisition.* Dordrecht: Reidel.

Fodor, J. D. 1989. Learning the periphery. In R. J. Matthews and W. Demopoulos (eds.),

Learnability and Linguistic Theory. Dordrecht: Kluwer. pp. 129-154.

Foley, W.A. and Van Valin, R.D. Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Fucci , D., Ellis, L. and Petrosino, L. 1990. Speech clarity/intelligibility: test-retest reliability of magnitude estimation scaling. *Perceptual and Motor Skills* 70: 232-234.

Gleitman, H. and Gleitman, L. 1979. Language use and language judgment. In C.J.Fillmore, W. Kempler, and W. S.-Y. Wang (eds.) , *Individual Differences in Language Ability and Language Beavior*. New York: Academic Press. pp. 103-126.

Gougenheim, G. 1951. *Grammaire de la langue française du seizième siècle*. Lyon: Editions I.A.C.

Green, K. 1987. The perception of speaking rate using visual information from a talker's face. *Perception and Psychophysics*. 42: 587-593.

Greenbaum, S. (ed.) 1977. *Acceptability in Language*. The Hague: Mouton.

Greenbaum, S. 1988. *Good English and the Grammarian*. London: Longman.

Greenbaum, S. and Quirk, R. 1970. *Elicitation experiments in English: Linguistic Studies in Use and Attitude*. London: Longmans.

Grevisse, M. 1980 (XI ed.). *Le bon usage. Grammaire française avec des remarques sur la langue française d'aujourd'hui*. Paris-Gembloux: Duculot.

Grewendorf, G. 1989. *Ergativity in German*. Dordrecht: Foris.

Grimshaw, J. 1987. Unaccusatives - An Overview. *Proceedings of the North-East Linguistic Society*, vol. 17: 244-258.

Grimshaw, J. 1991. *Argument Structure*. Cambridge, MA: MIT Press.

Gruber, J. 1965. *Studies in Lexical Relations*. PhD Dissertation, MIT.

Gruber, J. 1976. *Lexical Structures in Syntax and Semantics*. Amsterdam: North Holland.

Guéron, J. and Hoekstra, T. 1988. T-chains and the constituent structure of auxiliaries. In A.Cardinaletti, G. Cinque and G. Giusti (eds.), *Constituent Structure*. Dordrecht: Foris: 35-100.

Guerssel,M., Hale, K. Laughren, M. Levin, B. and White Eagle, J. 1985. A cross-linguistic study of transitivity alternations. *Papers from the Parasession on Causatives and Agentivity at the 21st Regional Meeting of the Chicago Linguistic Society*.

Haegeman, L. 1991. *Introduction to Government & Binding Theory*. Oxford: Blackwell.

Hale, K. 1988. Linguistic theory: generative grammar. In S. Flynn and W. O'Neill (eds.), *Linguistic Theory in Second Language Acquisition*. Dordrecht: Kluwer.

Hoekstra, T. 1984. *Transitivity: Grammatical Relations in Government-BindingTheory*. Dordrecht: Foris.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar.* Cambridge MA: MIT Press.

Jackendoff, R. 1983. *Semantics and Cognition.* Cambridge MA: MIT Press.

Jackendoff, R. 1991. *Semantic Structures.* Cambridge MA: MIT Press.

Jacqmain, M. and Meerts, E. 19??. Problemi di ausiliare. *Studi di grammatica italiana* 2: 215-244.

Kayne, R. 1975. *French Syntax.* Cambridge MA: MIT Press.

Kayne, R. 1980. **J'ai voulu lire, j'ai tout voulu lire. Langue Française* 46: 32-40.

Keenan, E. 1987. Semantic correlates of the ergative/absolutive distinction. In E.L. Keenan (ed.), *Universal Grammar: 15 Essays*. London: Croom Helm.

Keenan, E. (1987). Semantic correlates of the ergative/absolutive distinction. In E.L. Keenan (ed.), *Universal Grammar: 15 Essays*. London: Croom Helm.

Kellerman, E. 1978. Giving learners a break: Native language intuitions as a source of predictions about transferability. *Working Papers on Bilingualism* 15: 59-92.

Keyser, S.J. & Roeper, T. 1984. On the middle and ergative constructions in English. *Linguistic Inquiry* 15: 381-416.

Klein, W. 1986. *Second Language Acquisition.* Cambridge: Cambridge University Press.

Kohn, K. 1979. Aspekte lernsprachliche Verhaltens. Unpublished ms., University of Konstanz.

Kohn, K. 1982. Beyond output: The analysis of interlanguage development. *Studies in Second Language Acquisition* 4: 137-152.

Kroch, A. 1989. Language learning and language change. in D. Lightfoot, The child's trigger experience: degree-0 learnability. *Behavioral and Brain Sciences* 12: 321-375.

Krueger, L.E. 1989. Reconciling Fechner and Stevens: Toward a unified psychophysical law. *Behavioral and Brain Sciences* 12: 251-320.

Labelle, M. 1990. Unaccusatives and pseudo-unaccusatives in French. *NELS* 20: 303-317.

Labov, W. 1975. *What is a Linguistic Fact?* Lisse: Peter de Ridder Press.

Lakoff, G. 1973. Fuzzy grammar and the performance/competence terminology game. *Papers from the 9th Regional Meeting of the Chicago Linguistic Society* 271-291.

Legendre, G. 1989. Unaccusativity in French. *Lingua* 79: 95-164.

Levelt, W.J.M. 1974. *Formal Grammars in Linguistics and Psycholinguistics.* Vol. 3: Psycholinguistic applications. The Hague: Mouton.

Levelt, W.J.M., Haans, J.A.W.M., and Meijers, A.J.A. 1977. Grammaticality, paraphrase, and imagery. In Greenbaum, S. (ed.) *Acceptability in Language.* The Hague: Mouton.

Levin, L.S. 1985. *Operations on Lexical Form: Unaccusative Rules in Germanic Languages.* New York and London: Garland.

Liceras, J. 1983. *Markedness, Contrastive Analysis, and the Acquisition of Spanish Syntax by English Speakers.* PhD Thesis.

Lightfoot, D. 1979. *Principles of Diachronic Syntax.* Cambridge: Cambridge University Press.

Lightfoot, D. 1983. *The Language Lottery: Towards a Biology of Grammars*. Cambridge, MA: MIT Press.

Lightfoot, D. 1989. The child's trigger experience: Degree-0 learnability. *Behavioral and Brain Sciences* 12: 321-375.

Lightfoot, D. 1991. *How to Set Parameters: Arguments from Language Change.* Cambridge MA: MIT Press.

Lodge, M. 1981. *Magnitude Scaling. Quantitative Measurement of Opinions.* Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-025. Beverly Hills and London: Sage Publications.

Mereu, L. 1981. A contrastive study of verbal aspect in Italian and English. *ITL* 54: 3-26.

Miller, G. 1969. A psychological method to investigate verbal concepts. *Journal of Mathematical Psychology* 6: 169-191.

Mohan, B.A. 1977. Acceptability testing and fuzzy grammars. In Greenbaum, S. (ed.) *Acceptability in Language.* The Hague: Mouton.

Napoli, D.J. 1988. Review article on Burzio 1986. *Language* 64: 130-142.

Ouhalla, J. 1991. *Functional Categories and Parametric Variation.* London: Routledge.

Newmeyer, F.J. 1983. *Grammatical Theory: Its Limits and Its Possibilities.* Chicago: The University of Chicago Press.

Nunnally, J.C. 1967. *Psychometric Theory.* New York: McGraw-Hill.

Parisi, D. 1976. The past participle. *Italian Linguistics* 1: 77-106.

Pankhurst, J., van Buren, P. and Sharwood Smith, M. (eds). 1988. *Learnability and Second Languages*. Dordrecht: Foris.

Pateman, T. 1985. From nativism to sociolinguistics: Integrating a theory of language growth with a theory of speech practices. *Journal for the Theory of Social Behaviour* 15:38-59.

Pavlovic, C., Rossi, M. and Espesser, R. 1990. Use of magnitude estimation technique for assessing the performance of text-to-speech synthesis systems. *Journal of the Acoustical Society of America* 87: 373-382.

Pearce, E. 1990. *Parameters in Old French Syntax. Infinitival Complements*. Dordrecht: Foris.

Perlmutter, D. 1978. Impersonal passives and the Unaccusative Hypothesis. In *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: University of California.

Perlmutter, D. (ed.) 1983. *Studies in Relational Grammar 1*. Chicago: Chicago University Press.

Perlmutter, D. and P. Postal 1984. The 1-Advancement Exclusiveness Law. In D. Perlmutter and C. Rosen (eds.), *Studies in Relational Grammar 2*. Chicago: University of Chicago Press.

Perlmutter, D. 1989. Multiattachment and the unaccusative hypothesis: the perfect auxiliary in Italian. *Probus* 1: 63-119.

Pinker, S. 1989. Markedness and language development. In R. J. Matthews and W. Demopoulos (eds.), *Learnability and Linguistic Theory*. Dordrecht: Kluwer.

Pinker, S. 1991. *Learnability and Cognition. The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.

Pollock, J.-Y. 1985. Verb movement, universal grammar, and the structure of IP. *Linguistic inquiry* 20: 365-424.

Poulton, E. 1968. The new psychophysics: Six models for magnitude estimation. *Psychological Bulletin* 69: 1-19.

Poulton, E. 1968. The new psychophysics: six models for magnitude estimation. *Psychological Bullettin* 69: 1-19.

Poulton, E. 1989. *Bias in Quantifying Judgments*. Hove: Lawrence Erlbaum.

Pullum, G. K. 1991. *The Great Eskimo Vocabulary Hoax*. Chicago: University of Chicago Press.

Py, B. 1986. Native language attrition amongst migrant workers: towards an extension of the concept of interlanguage. In E. Kellerman and M. Sharwood Smith (eds.), *Crosslinguistic Influence in Second Language Acquisition*. Oxford: Pergamon.

Ramat, P. 1982. An example of reanalysis: periphrastic forms in the Romance languages' verb system. *Typological Linguistics*.

Rappaport, M. 1989. Unaccusativity and verbs of motion. In School of English, Aristotle University of Thessaloniki, Greece (ed.), *2nd Symposium on English and Greek: description and/or Comparison of the Two Languages*. Thessaloniki, Greece.

Rappaport, M. and Levin, B. 1988. What to do with theta-roles. In W.Wilkins (ed.), *Syntax and Semantics. Thematic Relations*. New York: Academic Press.

Ringen, D. 1974. Linguistic facts: A study of the empirical scientific status of transformatioal generative grammar. In D. Cohen and J.R.Wirth (eds.), *Testing Linguistic Hypotheses*. New York: John Wiley & Sons.

Rizzi, L. 1978. A restructuring rule in Italian syntax. In S.J. Keyser (ed.), *Recent Transformational Studies in European Languages*. Cambridge: MA: MIT Press.

Rizzi, L. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Roberts, I. 1990. Review of Zubizarreta 1987. *Natural Language and Linguistic Theory* 8: 465-489.

Rohlfs, G. 1969. *Grammatica storica della lingua italiana e dei suoi dialetti.* Torino: Einaudi.

Rosen, C. 1984. The interface between semantic roles and initial grammatical relations. In D. Perlmutter and C. Rosen (eds.), *Studies in Relational Grammar 2*. Chicago: Chicago University Press.

Ross, J.R. 1979. Where's English? In C.J.Fillmore, W. Kempler, and W. S.-Y. Wang (eds.) , *Individual Differences in Language Ability and Language Beavior.* New York: Academic Press. pp.127-165

Rydén, M. and Brorström, S. 1987. *The Be/Have Variation with Intransitives in English. With Special Reference to the Late Modern Period.* Acta Universitatis Stockholmiensis. Stockholm Studies in English LXX.

Salvi, G. 1987. Syntactic restructuring and the evolution of Romance auxiliaries. In M. Harris and P. Ramat (eds.) *Historical Development of Auxiliaries*. Amsterdam: Mouton de Gruyter.

Sankoff, G. and P. Thibault 1977. L'alternance entre les auxiliaires *avoir* et *être* en français parlé à Montréal. *Langue Française*: 81-108.

Savinian, A. 1882. *Grammair Provençale*. Avignon-Paris.

Schachter, J. 1989. Testing a proposed universal. In S. Gass and J. Schachter (eds.), *Linguistic Perspectives on Second Language Acquisition.* Cambridge: Cambridge University Press.

Schachter, J., Tyson, A. and Diffley, F. 1976. Learners' intuitions of grammaticality. *Language Learning* 26: 67-76.

Selinker, L. 1972. Interlanguage. *International Review of Applied Linguistics* 10.

Siegel, S. and Castellan, N.J. 1988. *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

Snow, C. and Meijer, G. 1977. On the secondary nature of syntactic intuitions. In Greenbaum, S. (ed.) *Acceptability in Language*. The Hague: Mouton.

Sorace, A. 1985a. Certainty and consistency in interlanguage intuitional data. Unpublished manuscript, University of Southern California.

Sorace, A. 1985b. Metalinguistic knowledge and language use in acquisition-poor environments. *Applied Linguistics* 6: 239-254.

Sorace, A. 1988. Linguistic intuitions in interlanguage development: the problem of indeterminacy. In J. Pankhurst, M. Sharwood Smith and P. van Buren (eds.), *Learnability and Second Languages. A Book of Readings*. Dordrecht: Foris.

Spencer, N.J. 1973. Differences between linguists and nonlinguists in intuitions of grammaticality-acceptability. *Journal of Psycholinguistic Research* 2: 83-98.

Sportiche, D. 1989. A theory of floating quantifiers. *Linguistic Inquiry* 19: 425-449.

Stevens, J.C., Mack, J.D. and Stevens, S.S. 1960. Growth of sensation on seven continua as measured by force of hadgrip. *Journal of Experimental Psychology*. 59: 60-67.

Stevens, S.S.1961. The psychophysics of sensory functions. In W.A. Rosenblith (ed.), *Sensory Communication*. Cambridge: Cambridge University Press.

Stevens, S.S. 1966. Ametric for the social consensus. *Science* 151: 530-541.

Stevens, S.S. 1971. Issues in psychophysical measurement. *Psychological Review*. 78:426-450.

Stevens, S.S. 1975. *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects*. New York: John Wiley.

Stowell, T. 1981. *Originas of Phrase Structure*. PhD Thesis, MIT.

Takefuta, Y., Guberina, P., Pizzamiglio, L., and Black, J. 1986. Cross-lingual measurement of interconsonantal differences. *Journal of Psycholinguistioc Research* 15: 489-507.

Toner, M . and Emmanuel, F. 1989. Direct magnitude estimation and equal appearing scaling of vowel roughness.*Journal of Speech and Hearing Research*. 32: 78-82.

Traugott, E.C., and Heine, B. (eds.) 1991. *Approaches to Grammaticalization*. Amsterdam: John Benjamins.

Tsimpli, M-I and Roussou, A, 1991. Parameter--resetting in SLA? *University College London Working Papers in Linguistics* 3.

Tuttle, D. 1986. The spread of *esse* as universal auxiliary in central Italo-Romance. *Medioevo Romanzo* XI.

Van Valin, R.D. 1990. Semantic parameters of split intransitivity. *Language* 66: 221-260.

Vendler, Z. 1967. *Linguistics in Philosophy*. Ithaca: Cornell University Press.

Vincent, N. 1982. The development of the auxiliaries *habere* and *esse* in Romance. In N. Vincent and M. Harris (eds.), *Studies in the Romance Verb*. London: Croom Helm.

Vincent, N. 1987. Italian. In B. Comrie (ed.), *The World's Major Languages*. London: Croom Helm. pp. 279-302.

Vincent, N. 1989. Observing obsolescence. In D. Lightfoot, The child's trigger experience: degree-0 learnability. *Behavioral and Brian Sciences* 12: 321-375.

White, L. 1985. The 'pro-drop' parameter in adult second language acquisition. *Language Learning* 35: 47-62.

White, L. 1988. Island effects in second language acquisition. In S. Flynn and W. O'Neil (eds.), *Linguistic Theory in Second Language Acquisition*. Dordrecht: Kluwer.

White, L. 1989. *Universal Grammar and Second Language Acquisition.* Amsterdam: John Benjamins.

White, L. forthcoming. Universal Grammar and second language acquisition: Current trends and new directions. To appear in W. Ritchie and T. Bhatia (eds.), *Handbook of Language Acquisition.* New York: Academic Press.

Williams, E. 1981. Language acquisition, markedness, and phrase structure. In S. Tavakolian (ed.), *Language Acquisition and Linguistic Theory.* Cambridge MA: MIT Press.

Wode, H. 1978. Developmental sequences in naturalistic L2 acquisition. In E.M. Hatch (ed.), *Second Language Acquisition*. Rowley, MA: Newbury House.

Yule, G., Yanz, J.L., and Tsuda, A. 1985. Investigating aspects of the language learner's confidence: an application of the theory of signal detection. *Language Learning*

Zobl, H. (1989). Canonical typological structures and ergativity in English L2 acquisition. In S. Gass and J. Schachter (eds.), *Linguistic Perspectives on Second Language Acquisition*. Cambridge: Cambridge University Press, 203-221.

Zubizarreta, M-L 1987. *Levels of Representation in the Lexicon and in the Syntax.* Dordrecht: Foris.

APPENDIX A

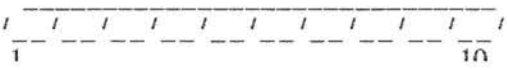
TEST INSTRUCTIONS AND SENTENCES

A.1 PILOT 1

INSTRUCTIONS TO ITALIAN LEARNERS OF FRENCH

Lo scopo di questa ricerca è l'analisi delle **intuizioni linguistiche** sviluppate durante l'apprendimento di una lingua straniera. Questo test è completamente anonimo.

Avete di fronte a voi una lista di 58 frasi. Dovrete giudicare l'**accettabilità grammaticale** di queste frasi: chiedetevi se la frase "suona" bene oppure no, senza cercare di ricordare regole grammaticali. Le vostre reazioni immediate sono l'obiettivo principale di questo test. Sotto ogni frase sono riportate delle caselle numerate da 1 a 10:

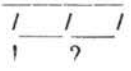


Questi numeri significano:

1 = la frase è perfettamente accettabile
10 = la frase completamente inaccettabile

Alcune frasi sono più (o meno) accettabili di altre. Potete usare le caselle tra 1 e 10 per le frasi che, secondo voi, hanno un grado di accettabilità "intermedio".

Sotto ogni frase troverete anche due caselle contrassegnate da un punto eslamativo e da un punto interrogativo:



Queste caselle servono ad esprimere il grado di **certezza** dei vostri giudizi sull'accettabilità delle frasi e significano:

! = certezza ? = incertezza

Grazie della vostra collaborazione!

PILOT 1

INSTRUCTIONS TO FRENCH LEARNERS OF ITALIAN

Le but de cette recherche est l'analyse des **intuitions linguistiques** qui se développent pendant l'apprentissage d'une langue étrangère.
Ce test est complètement anonyme.

Vous avez devant vous une liste de 58 phrases. Vous devez juger **l'acceptabilité grammaticale** de ces phrases: posez-vous la question si chaque phrase vous semble "bonne" ou pas, sans essayer de vous rappeler des règles grammaticales. Vos réactions immédiates sont l'objectif principal de ce test.

Sous chaque phrase se trovent des cases numérotées de 1 à 10. Ces chiffres signifient:

1 = phrase parfaitement acceptable
10 = phrase complètement inacceptable

Quelques phrases sont plus (ou moins) acceptables que d'autres. Vous pouvez utilizer les cases entre 1 et 10 pour les phrases qui, selon vous, ont un degré d'acceptabilité intermédiaire.

Sous chaque phrase vous trouvez aussi deux cases marquées par un point d'exclamation et un point d'interrogation:

!		?

Ces cases servent à exprimer le degré de **certitude** de vos jugements sur l'acceptabilité des phrases. Elles signifient:

1 = je suis sûr de mon jugement
? = je ne suis pas sûr de mon jugement

Merci de votre collaboration!

PILOT 1

TEST SENTENCES

(a) Sentences judged by French learners of Italian : correct auxiliaries

UNERGATIVE / MOTIONAL

Mia sorella ha nuotato nella piscina
Abbiamo camminato per tre ore
Michela ha viaggiato molto da giovane

UNERGATIVE / NON-MOTIONAL

Carla ha dormito fino a mezzogiorno
I ragazzi hanno giocato tutto il pomeriggio
Cristina ha lavorato a Londra per un anno

UNERGATIVE / [+UNACCUSATIVE ALTERNANT]

Paola ha corso più velocemente di Maria
I bambini hanno saltato sul letto per due ore
La palla ha rotolato lungo il corridoio

UNACCUSATIVE / CHANGE OF STATE (I = F)

Le foglie sono cadute dagli alberi
Paola è ritornata a casa alle nove
Mia cugina è diventata ricca in America

UNACCUSATIVE / CHANGE OF STATE (I ≠ F)

I soldi sono spariti dal cassetto
I miei nonni sono invecchiati molto l'anno scorso
La bambina è arrossita dalla vergogna

UNACCUSATIVE / CONTINUATION OF STATE

Le mie amiche sono restate da me fino a tardi
I miei nonni sono sopravvissuti al terremoto del 908

UNACCUSATIVE / EXISTENCE OF STATE

I dinosauri sono esistiti milioni di anni fa
Quella storia mi è sembrata incredibile
Questa pipa è appartenuta a mio nonno

UNACCUSATIVE / [+TRANSITIVE ALTERNANT]

Le tasse sono aumentate del 20%
La nave è affondata rapidamente
La casa dei miei genitori non è cambiata dalla mia infanzia

UNACCUSATIVE / [INTRANSITIVE ALTERNANT]

La palla è rotolata nel buco
I bambini sono saltati giù dal letto
Paola è corsa in farmacia

RAISING CONSTRUCTIONS

Maria è sembrata risolvere il problema
Paola e sua figlia sono dovute partire alle otto
Maria è potuta rimanere a Firenze tutta l'estate

(a) Sentences judged by French learners of Italian : incorrect auxiliaries

UNERGATIVE / MOTIONAL

Mia figlia è nuotata tutti i giorni quest'estate
Siamo camminati fino al centro della città
Paola è viaggiata molto negli Stati Uniti

UNERGATIVE / NON-MOTIONAL

Carla è dormita fino alle dieci
Dopo cena gli ospiti sono giocati a carte
Giovanna è lavorata a Perugia per due anni

UNACCUSATIVE / CHANGE OF STATE (I = F)

I bambini hanno caduto nella piscina
Carla ha ritornato al suo paese dopo vent'anni
Angela ha diventato un'attrice famosa

UNACCUSATIVE / CHANGE OF STATE (I ≠ F)

Molti libri hanno sparito dalla biblioteca
Mia nonna non ha invecchiato molto
Maria ha arrossito dall"imbarazzo

UNACCUSATIVE/CONTINUATIVE

Luisa e Lina hanno restato in montagna tutta l'estate
Molti argentini hanno sopravvissuto alle persecuzioni politiche

UNACCUSATIVE/EXISTENTIAL

Gli unicorni non hanno mai esistito
La commedia mi ha sembrato noiosa
Questa penna ha appartenuto ai miei bisnonni

UNACCUSATIVE WITH TRANSITIVE ALTERNANT

I prezzi hanno aumentato del 0% il mese scorso
Le navi hanno affondato a tre Km dal porto
La casa dei miei nonni ha cambiato molto recentemente

UNACCUSATIVE WITH INTRANSITIVE ALTERNANT

La moneta ha rotolato giù dal tavolo
La moneta è rotolata velocemente
Maria ha corso a casa
Maria è corsa meno velocemente di Paola
L'atleta è saltata 2m50 alle Olimpiadi
I ladri hanno saltato giù dalla finestra

RAISING

Paola ha sembrato capire la situazione
Maria e Giovanna hanno dovuto restare a casa
Maria non ha potuto venire alla festa

(b) Sentences judged by Italian learners of French : correct auxiliaries

UNERGATIVE / MOTIONAL

Ma soeur a nagé dans la piscine
Nous avons marché pendant trois heures
Ma tante a beaucoup voyagé pendant sa jeunesse

UNERGATIVE / NON-MOTIONAL

Chantal a dormi jusqu'à midi
Les enfants ont joué pendant tout l'après-midi
Christine a travaillé à Londres pendant un an

UNERGATIVE / [+UNACCUSATIVE ALTERNANT]

Pauline a couru plus vite que Marie-France
Les enfants ont sauté sur le lit des parents toute la matinée
Les enfants ont sauté du lit a sept heures

UNACCUSATIVE / CHANGE OF STATE (I = F)

Les feuilles sont tombées des arbres
Pauline est rentrée chez elle à neuf heures
Ma cousine est devenue très riche aux Etats-Unis

UNACCUSATIVE / CHANGE OF STATE (I ≠ F)

Les documents ont disparu de ma serviette
Mes grands-parents ont beaucoup vieilli cet hiver
La petite fille a rougi de honte

UNACCUSATIVE / CONTINUATION OF STATE

Mes amies sont restées très tard chez moi
Mes grands-parents ont survécu au tremblement de terre de 1908

UNACCUSATIVE / EXISTENCE OF STATE

Les dinosaures ont existé il y a plusieurs millions d'années
Cette histoire m'a semblé invraisemblable
Cette pipe a appartenu à mon grand-père

UNACCUSATIVE / [+TRANSITIVE ALTERNANT]

Les impôts ont augmenté de 20%
Les bateaux ont coulé très vite
La maison de mes parents a très peu changé depuis mon enfance

UNACCUSATIVE / [+UNERGATIVE ALTERNANT]

La bille a roulé dans le trou
La bille a roulé le long du couloir
Pauline a couru à la pharmacie

RAISING

Marie-Claude a semblé résoudre le problème
Pauline et sa fille ont dû partir à huit heures
Marie-Françoise a pu rester à Florence tout l'été

(b) Sentences judged by Italian learners of French : incorrect auxiliaries

UNERGATIVE / MOTIONAL

Ma fille est nagée tous les jours l'été dernier
Nous sommes marchés jusqu'au centre de la ville
Pauline est beaucoup voyagée aux Etats-Units

UNERGATIVE / NON-MOTIONAL

Chantal est dormie jusqu'à dix heures
Après le dinêr, les invités sont joués aux cartes
Jeanne est travaillée à Paris pendant deux ans

UNERGATIVE / [+UNACCUSATIVE ALTERNANT]

Pauline est courue moins vite que Caroline
L'athlète est sautée50 aux Jeux Olimpiques
Les cambrioleurs sont sautés par la fenêtre

UNACCUSATIVE / CHANGE OF STATE (I = F)

Les enfants ont tombé dans la piscine
Caroline a rentré dans son pays au bout de 20 ans
Marianne a devenu une actrice célèbre

UNACCUSATIVE / CHANGE OF STATE (I = F)

Plusieurs livres sont disparus de la bibliothèque
Ma grand'mère n'est pas beaucoup vieillie
Arlette est rougie de gêne

UNACCUSATIVE / CONTINUATION OF STATE

Jeanne et Mireille ont resté à la montagne tout l'été
Beaucoup d'Argentins sont survécus aux persecutions politiques

UNACCUSATIVE / EXISTENCE OF STATE

Les unicornes ne sont jamais existés
La pièce m'est semblée ennuyeuse
Cette chaise est appartenue à mes ancêtres

UNACCUSATIVE / [+TRANSITIVE ALTERNANT]

Les prix sont augmentés de 0% le mois dernier
Les bateaux sont coulés à 3 Km du port
La maison de mes parents est beaucoup changée récemment

UNACCUSATIVE / [+UNERGATIVE ALTERNANT]

La pièce de monnaie est roulée de sur la table
La pièce de monnaie est roulée très vite
Pauline est courue à la maison

RAISING

Pauline est semblée comprendre la situation
Chantal et Véronique sont dûes rester à la maison
Véronique n'est pas pue venir à la fête

A.2 PILOT 2

PILOTS 2 AND 3
INSTRUCTIONS TO NATIVE ITALIAN SPEAKERS

Lo scopo di questo esperimento è verificare come i parlanti di madre lingua italiana giudicano l'accettabilità di alcune strutture dell'italiano e, in particolare, come ordinano certe frasi l'una rispetto all'altra. Nel rispondere a questi test, dovreste chiedervi se una certa frase 'suona' più accettabile di un'altra. Cercate, nei limiti del possibile, di distinguere tra vari 'gradi' di accettabilità (o di inaccettabilità).

Magnitude estimation
Vedrete delle frasi isolate, una alla volta, proiettate sullo schermo. Dovrete giudicare l'accettabilità di queste frasi. Date alla prima frase un numero a caso. Poi, per ogni frase successiva, date un numero che sia proporzionale al primo numero. Per esempio, se avete dato alla prima frase il numero 6 e pensate che l'accettabilità della seconda frase sia un terzo della prima, date il numero 2 a questa seconda frase. Se invece pensate che sia due volte più accettabile della prima, datele il numero 12. Se pensate che la seconda frase sia leggermente meno accettabile della prima, usate il numero 5, e così via.

Cards
Avete una serie di striscie di carta, ciascuna rappresentante una frase. Cercate di raggruppare le frasi a seconda del loro grado di accettabilità: potete formare quanti gruppi volete. Poi sistemate i gruppi di frasi lungo l'unità di misura di fronte a voi, partendo dalle frasi meno accettabili sulla sinistra fino alle frasi più accettabili sulla destra. Per ogni posizione lungo l'unità di misura, segnate i numeri corrispondenti alle frasi.

Rank order
Vedete di fronte a voi una lista di 24 frasi. Alcune frasi sono perfettamente accettabili e normali, altre lo sono di meno. Dovete formare una graduatoria, mettendo le frasi in ordine di accettabilità, secondo questo criterio:
1 = perfettamente accettabile 24 = completamente inaccettabile.
Potete dare lo stesso numero a più di una frase, se pensate che abbiano lo stesso grado di accettabilità.

PILOTS 2 AND 3
INSTRUCTIONS TO ENGLISH LEARNERS OF ITALIAN

The aim of this experiment is to investigate how learners of Italian judge the acceptability of some Italian structures and particularly how they rank order sentences. In doing these tests, ask yourselves if a certain sentence 'sounds' better than another. Please try to distinguish as many degrees of acceptability (or unacceptability) as you can.

Magnitude estimation

Isolated sentences will be projected on the screen in front of you, one at a time. You will have to judge the acceptability of these sentences. Give the first sentence any number you wish; then assign successive sentences numbers that are proportional to the first number you chose. For instance, if you gave the number 6 to the first sentences and if you think that the second sentence is one third as acceptable as the first, give the number 2 to it. Similarly, if you think that the second sentence is twice as acceptable than the first one, choose the number 12. If you think that the second sentence is only a little less acceptable than the first, use the number 5, and so on.

Cards

You have a set of cards, each representing a sentence. Your task is to arrange sentences in different piles according to their degree of acceptability. You may form as many piles as you wish. Once you have divided all the sentences into piles, place the piles along the tape measure in front of you, starting from 'least acceptable' on your left towards 'most acceptable' on your right. For each position you choose along the tape measure, please write the numbers corresponding to the sentences in that pile.

Rank order

You have a list of 24 (30) sentences in front of you. Some of them are perfectly acceptable and normal, others are less acceptable. Your task is to rank the sentences according to their degree of acceptability, by giving them numbers from 1 to 30 so that:

1 = least acceptable 24 (30) = most acceptable

You may give the same number to more than one sentence if you think they have the same degree of acceptability.

PILOT 2
LEXICALIZATION 1 / GRAMMATICAL SENTENCES

(1) Maria è andata in ufficio a piedi
(2) Le mie amiche sono rimaste a casa mia fino a tardi
(3) Questa pipa è appartenuta a mio nonno
(4) Paolo ha baciato Luisa appassionatamente
(5) Maria ha lavato tutte le finestre
(6) L'anno scorso il governo ha aumentato le tasse
(7) Carla ha passeggiato nel parco tutto il pomeriggio
(8) Maria ha dormito fino a mezzogiorno
(9) Paola è corsa a casa di Piero
(10) Simona ha corso più velocemente di Maria
(11) Mia figlia è migliorata molto a scuola
(12) Cristina ha vissuto a Parigi dal 1981 al 1985

LEXICALIZATION 1 / UNGRAMMATICAL SENTENCES

(13) Maria ha andato in ufficio a piedi
(14) Le mie amiche hanno rimasto da me fino a tardi
(15) Questa pipa ha appartenuto a mio nonno
(16) Paolo ha baciata Luisa appassionatamente
(17) Maria ha lavate tutte le finestre
(18) L'anno scorso il governo ha aumentate le tasse
(19) Carla è passeggiata nel parco tutto il pomeriggio
(20) Maria è dormita fino a mezzogiorno
(21) Paola ha corso a casa di Piero
(22) Simona è corsa più velocemente di Maria
(23) Mia figlia ha migliorato molto a scuola
(24) Cristina è vissuta a Parigi dal 1981 al 1985

PILOT 2
LEXICALIZATION 2 / GRAMMATICAL SENTENCES

(1) Paola è uscita di casa alle nove
(2) Anna è restata a casa degli zii tutta l'estate
(3) La commedia è piaciuta molto al pubblico
(4) Mauro ha suonato la chitarra alla festa
(5) Il gatto ha rotto tre bicchieri
(6) Luigi ha cambiato i soldi all'aeroporto
(7) Mia sorella ha nuotato nella piscina olimpionica
(8) I ragazzi hanno giocato tutto il pomeriggio
(9) La palla è rotolata giù nel fossato
(10) La palla ha rotolato lungo il corridoio
(11) La lezione di matematica è cominciata alle dieci
(12) I due aerei sono decollati contemporaneamente

LEXICALIZATION 2 / UNGRAMMATICAL SENTENCES

(13) Paola ha uscito di casa alle nove
(14) Anna ha restato a casa degli zii tutta l'estate
(15) La commedia ha piaciuto molto al pubblico
(16) Mauro ha suonata la chitarra alla festa
(17) Il gatto ha rotti tre bicchieri
(18) Luigi ha cambiati i soldi all'aeroporto
(19) Mia sorella è nuotata nella piscina olimpionica
(20) I ragazzi sono giocati tutto il pomeriggio
(21) La palla ha rotolato giù nel fossato
(22) La palla è rotolata lungo il corridoio
(23) La lezione di matematica ha cominciato alle dieci
(24) I due aerei hanno decollato contemporaneamente

PILOT 2
LEXICALIZATION 3 / GRAMMATICAL SENTENCES

(1) I miei genitori sono arrivati in macchina
(2) I miei nonni sono sopravvissuti al terremoto del 1908
(3) La storia non è sembrata credibile alla polizia
(4) Il padrone di casa ha stretto la mano all'ospite
(5) Giovanni ha mangiato la mia tavoletta di cioccolato
(6) Il direttore ha bruciato le lettere compromettenti
(7) Michela ha viaggiato molto da giovane
(8) Giovanna ha lavorato a Pisa per un anno
(9) Quando mi sono avvicinato, i serpenti sono strisciati via
(10) I serpenti hanno strisciato lentamente verso la preda
(11) L'anno scorso i prezzi sono aumentati del 10%
(12) Ieri ha piovuto tutto il giorno

LEXICALIZATION 2 / UNGRAMMATICAL SENTENCES

(13) I miei genitori hanno arrivato in macchina
(14) I miei nonni hanno sopravvissuto al terremoto del 1908
(15) La storia non ha sembrato credibile alla polizia
(16) Il padrone di casa ha stretta la mano all'ospite
(17) Giovanni ha mangiata la mia tavoletta di cioccolato
(18) Il direttore ha bruciate le lettere compromettenti
(19) Michela è viaggiata molto da giovane
(20) Giovanna è lavorata a Pisa per un anno
(21) Quando mi sono avvicinato, i serpenti hanno strisciato via
(22) I serpenti sono strisciati lentamente verso la preda
(23) L'anno scorso i prezzi hanno aumentato del 10%
(24) Ieri è piovuto tutto il giorno

A.3 PILOT 3

LEXICALIZATION 1 / GRAMMATICAL SENTENCES

(1) Ieri a quest'ora dormivo
(2) Quando ha telefonato Maria, guardavo la TV
(3) Aspettavo l'autobus quando ho visto Paolo
(4) Guidavo dalle 6 di mattina quando ho avuto l'incidente
(5) Ieri alle due distribuivo manifesti in via Cavour
(6) Dopo la passeggiata avevamo molta fame
(7) Dovevo partire, ma ho perso il treno
(8) Da giovane Mario portava i capelli lunghi
(9) In quel periodo Giovanna viveva con i nonni
(10) L'estate scorsa andavamo al mare tutti i fine-settimana
(11) Ho avuto una casa al mare per due anni
(12) Ieri ho dormito fino a mezzogiorno
(13) Il gatto ha rotto un bicchiere
(14) Stamattina ho telefonato a Mario due volte
(15) Per sei anni ho studiato il pianoforte tutti i giorni

LEXICALIZATION 1 / UNGRAMMATICAL SENTENCES

(1) Ieri a quest'ora ho dormito
(2) Quando ha telefonato Maria, ho guardato la TV
(3) Ho aspettato l'autobus da dieci minuti quando ho visto Franca
(4) Ho guidato dalle sei di mattina quando ho avuto l'incidente
(5) Ieri alle due ho distribuito manifesti in via Cavour
(6) Dopo la passeggiata abbiamo avuto molta fame
(7) Sono dovuto partire, ma ho perso il treno
(8) Da giovane Mario ha portato i capelli lunghi
(9) In quel periodo Giovanna ha vissuto con i nonni
(10) L'estate scorsa siamo andati al mare tutti i fine-settimana
(11) Avevo una casa al mare per due anni
(12) Ieri dormivo fino a mezzogiorno
(13) Il gatto rompeva un bicchiere
(14) Ieri telefonavo a Mario due volte
(15) Per tre anni studiavo il pianoforte tutti i giorni

PILOT 3
LEXICALIZATION 2 / GRAMMATICAL SENTENCES

(1) Ieri alle due ascoltavo musica classica
(2) Quando ha suonato il telefono, leggevo il giornale
(3) Passeggiavo da due ore quando ho incontrato Maria
(4) Vivevo in Italia dal 1981 quando ho sposato Paolo
(5) Ieri alle 5 scrivevo delle lettere
(6) Ho preso un taxi perché avevo paura di arrivare in ritardo
(7) Dovevo telefonare a Paolo, ma non avevo il numero
(8) Da bambina mia sorella portava gli occhiali
(9) In quei giorni studiavo in biblioteca
(10) L'estate scorsa giocavo a tennis tutti i venerdì
(11) Stamattina ho aspettato l'autobus per mezz'ora
(12) Ieri abbiamo guardato la TV fino a mezzanotte
(13) Stamattina la sveglia ha suonato alle sette
(14) Mio fratello e venuto a trovarmi molte volte
(15) Per tre ore ho telefonato a Maria ogni cinque minuti

LEXICALIZATION 2 / UNGRAMMATICAL SENTENCES

(1) Ieri alle due ho ascoltato musica classica
(2) Quando ha suonato il telefono, ho letto il giornale
(3) Ho passeggiato da due ore quando ho incontrato Maria
(4) Ho vissuto in Italia dal 1981 quando ho sposato Paolo
(5) Ieri alle cinque ho scritto delle lettere
(6) Ho preso un taxi perché ho avuto paura di arrivare in ritardo
(7) Ho dovuto telefonare a Maria, ma non avevo il numero
(8) Da bambina mia sorella ha portato gli occhiali
(9) In quei giorni ho studiato in biblioteca
(10) L'estate scorsa abbiamo giocato a tennis tutti i venerdì
(11) Stamattina aspettavo l'autobus per mezz'ora
(12) Ieri guardavamo la TV fino a mezzanotte
(13) Stamattina la sveglia suonava alle sette
(14) Mio fratello veniva a trovarmi molte volte
(15) Per tre ore telefonavo a Maria ogni cinque minuti

LEXICALIZATION 3 / GRAMMATICAL SENTENCES

(1) Ieri alle sette parlavo al telefono
(2) Quando ho bussato alla porta, Paolo dormiva
(3) Dormivo da due ore quando hai telefonato
(4) Lavoravo dalle cinque di mattina quando sei arrivato
(5) Ieri alle cinque facevo fotografie al Colosseo
(6) Mario è andato a casa presto perche aveva sonno
(7) Dovevo alzarmi presto ma non ho sentito la sveglia
(8) Da piccolo i pomodori non mi piacevano
(9) L'inverno scorso studiavo con Paolo
(10) L'inverno scorso andavamo spesso a sciare
(11) Ho studiato all'università per quattro anni
(12) Siamo rimasti al ristorante fino all'una
(13) All'improvviso c'è stata un'esplosione
(14) Mario è andato all'estero solo due volte
(15) Per due mesi sono andato a lezione d'italiano tutti i giorni

LEXICALIZATION 3 / UNGRAMMATICAL SENTENCES

(1) Ieri alle sette ho parlato al telefono
(2) Quando ho bussato alla porta, Paolo ha dormito
(3) Ho dormito da due ore quando hai telefonato
(4) Ho lavorato dalle cinque di mattina quando sei arrivato
(5) Ieri alle cinque ho fatto fotografie al Colosseo
(6) Mario è andato a casa presto perche ha avuto sonno
(7) Ho dovuto alzarmi presto ma non ho sentito la sveglia
(8) Da piccolo i pomodori non mi sono piaciuti
(9) L'inverno scorso ho studiato con Paolo
(10) L'inverno scorso siamo andati spesso a sciare
(11) Studiavo all'università per quattro anni
(12) Rimanevamo al ristorante fino all'una
(13) All'improvviso c'era un'esplosione
(14) Mario andava all'estero solo due volte
(15) Per due mesi andavo a lezione d'italiano tutti i giorni

A.4 MAIN STUDY

INSTRUCTIONS TO NATIVE ITALIAN SPEAKERS

Lo scopo di questo esperimento è verificare come i parlanti di madre lingua italiana giudicano l'accettabilità di alcune strutture dell'italiano e, in particolare, come ordinano certe frasi l'una rispetto all'altra. Nel rispondere a questi test, chiedetevi se una certa frase 'suona' più accettabile di un'altra. Vi accorgerete che alcune frasi sono completamente accettabili o inaccettabili (e provocheranno la vostra reazione immediata), mentre altre lo sono ad un livello intermedio (e avrete qualche dubbio nel decidere). Cercate, nei limiti del possibile, di distinguere tra vari 'gradi' di accettabilità (o di inaccettabilità). Fidatevi della vostra prima impressione, senza cercare di ricordare regole grammaticali.

L'esperimento consiste nei due test A e B. Alcuni di voi risponderanno prima al Test A e poi al Test B, altri prima al Test B e poi al Test A. Tutti i materiali di cui avrete bisogno si trovano nella vostra cartella.

Test A

Vedrete delle frasi isolate, una alla volta, proiettate sullo schermo. Dovrete giudicare l'accettabilità di queste frasi. Assegnate alla prima frase un numero a caso (diverso da zero) e trascrivetelo sul foglio delle risposte. Poi, per ogni frase successiva, date un numero che sia proporzionale al primo numero. Per esempio, se avete dato alla prima frase il numero 6 e pensate che la seconda frase sia due volte più accettabile della prima, date il numero 12 a questa seconda frase; se invece pensate che il suo grado di accettabilità sia un terzo della prima, date il numero 2; se pensate che sia leggermente meno accettabile della prima, date il numero 5, e così via. Potete assegnare lo stesso numero a frasi che, secondo voi, hanno lo stesso grado di accettabilità.

Esempio

Supponete che la prima frase sia

(1) Maria ha mangiato gli spaghetti

Questa frase è accettabile. Assegnatele, per esempio, il numero 12. La seconda frase potrebbe essere:

(2) Maria ha mangiati gli spaghetti

Questa frase, per alcuni, può avere un valore di accettabilità corrispondente alla metà della prima. Il numero assegnato è quindi 6. Supponete infine che la terza frase sia:

(3) Maria è gli spaghetti mangiato

L'accettabilità di questa frase potrebbe essere, per esempio, un terzo della prima: così scriverete il numero 4.

Ricordate: più grande è il numero, maggiore è l'accettabilità della frase.

Test B

Avete una serie di striscie di carta, ciascuna con una frase. Cercate di raggruppare le frasi a seconda del loro grado di accettabilità: potete formare quanti gruppi volete, anche di una sola frase. Una volta che avrete formato i gruppi, sistemateli sul tavolo, partendo dalle frasi meno accettabili sulla sinistra fino alle frasi più accettabili sulla destra. Poi usate il foglio di carta millimetrata per riprodurre l'ordine dei gruppi che avete formato. Trascrivete in colonna i numeri corrispondenti alle frasi contenute in ciascun gruppo, da sinistra (= frasi meno accettabili) verso destra (= frasi più accettabili).

Esempi

Poniamo che sette delle frasi siano queste:

(1) Maria ha mangiato gli spaghetti
(2) Maria è gli spaghetti mangiato
(3) E' mangiare gli spaghetti che è piaciuto a Maria
(4) Di spaghetti, Maria ha mangiati molti
(5) Di spaghetti, Maria ne ha mangiati molti
(6) E' mangiati gli spaghetti che sono piaciuti a Maria
(7) Maria ha mangiati gli spaghetti

Le frasi (1), (3), (5) sono tutte accettabili, e quindi possono essere messe nello stesso gruppo. Le frasi (2) e (6) sono completamente inaccettabili, quindi appartengono allo stesso gruppo. Le frasi (4) e (7) sono meno accettabili di quelle del primo gruppo, ma più accettabili di quelle del secondo gruppo: quindi formano un gruppo a parte (però potreste decidere che (4) è meno accettabile di (7): in questo caso, le due frasi formano due gruppi distinti).

Una volta formati questi gruppi, trascriveteli sul foglio di carta millimetrata (incluso nella cartella), come nel seguente esempio:

2	4	1
6	7	3
		5

MAIN STUDY
INSTRUCTIONS TO ENGLISH LEARNERS OF ITALIAN

The aim of this experiment is to see how foreign speakers of Italian judge the acceptability of some Italian structures and particularly how they rank order sentences with respect to one another. In answering these tests, ask yourself whether any given sentence "sounds" more or less acceptable than another. You will soon realize that certain sentences are completely acceptable or unacceptable (and you will have an immediate reaction to them), whereas other sentences will be (un)acceptable to a degree (and it will probably take you longer to decide). Please try to distinguish as many degrees of acceptability as you can. Trust your first impression, without trying to remember grammatical rules.

The experiment consists of two tests, A and B.
For Test A you are provided with a numbered answer sheet.
For Test B you are given a sheet of graph paper and an envelope containing printed sentences.

Some of you will do Test A first and then Test B, others will do Test B first and then Test A. All the materials for both tests are in your folder.

Thank you for your collaboration!

Test A

Isolated sentences will be projected on the screen in front of you, one at a time. You will have to judge the acceptability of these sentences. Give the first sentence any number you wish (but not 0); then assign successive sentences numbers that are proportional to the first number you chose. For instance, if you gave the number 6 to the first sentence and if you think that the second sentence is twice as acceptable as the first, give it the number 12. Similarly, if you think that the second sentence is only a third as acceptable as the first one, choose the number 2. If the second sentence seems to you only a little less acceptable than the first, give the number 5, and so on. You can assign the same number to different sentences if you think that they have the same degree of acceptability.

Examples

Suppose that the first sentence you see is the following:

(1) Maria ha mangiato gli spaghetti

This is an acceptable sentence and you give it, say, number 10. Then you see the second sentence:

(2) Maria ha mangiati gli spaghetti

You may think that this sentence is half as acceptable as the first one: you give it 5. The third sentence is:

(3) Maria è gli spaghetti mangiato

This sentence may look even less acceptable, and you assign it the number 3.

Remember: the higher the number, the more acceptable the sentence.

Test B

You have a set of cards, each with a sentence printed on it. Your task is to arrange the sentences in different piles according to their degree of acceptability. You may form as many piles as you wish, even if a pile contains only one sentence. Once you have divided all the sentences into piles, arrange the piles across your desk, starting from 'least acceptable' on your left towards 'most acceptable' on your right. Then use the sheet of graph paper to record which sentences are in which pile and how the piles are arranged across the desk. Write columns of numbers corresponding to the sentences in each pile, again from left (= least acceptable) to right (= most acceptable).

Examples

Suppose that seven of the cards have the following sentences printed on them:

(1) Maria ha mangiato gli spaghetti
(2) Maria è gli spaghetti mangiato
(3) E' mangiare gli spaghetti che è piaciuto a Maria
(4) Di spaghetti, Maria ha mangiato molti
(5) Di spaghetti, Maria ne ha mangiati molti
(6) E' mangiati gli spaghetti che sono piaciuti a Maria
(7) Maria ha mangiati gli spaghetti

Sentences (1), (3) and (5) are all acceptable and can therefore be assigned to the same group.
Sentences (2) and (6) are completely unacceptable, therefore they belong to a different group.
Sentences (4) and (7) are less acceptable than those belonging to the first group, but more acceptable than those of the second group: they can be put in a separate pile (of course, if you think that (4) is less acceptable than (7), then you should put the two sentences in two separate piles!).

Your groups are arranged as follows:	2	4	1
	6	7	3
			5

Finally, sentence numbers are written on the graph paper in the same order:

MAIN STUDY
INSTRUCTIONS TO FRENCH LEARNERS OF ITALIAN

Le but de cette recherche est d'analyser comment les Français qui parlent l'italien comme langue étrangère jugent l'acceptabilité de certaines structures de l'italien, et surtout comment ils ordonnent des phrases par rapport l'une à l'autre. Vous devrez vous poser la question si chaque phrase vous semble plus ou moins acceptable que les autres. Vous vous rendrez bientôt compte que certaines de ces phrases sont complètement acceptables ou inacceptables (et vous aurez une réaction immédiate), tandis que d'autres phrases ont un dégré d'acceptabilité intérmediare (et il vous faudra plus de temps pour décider). Essayez de distinguer autant de dégrés d'acceptabilité que possible. Vos réactions immédiates sont l'objectif principal de ce test: n'essayez pas de vous rappeler les règles grammaticales.

Le test comprend deux parties, A et B.
Pour le test A, vous utiliserez une feuille de réponses numérotée.
Pour le test B, vous aurez une feuille de papier millimétré et une enveloppe contenant des phrases imprimées.

Certains d'entre vous commenceront par le test A, tandis que d'autres commenceront par le test B. Tous les matériaux pour les deux tests se trouvent dans votre dossier.

Merci de votre collaboration!

Test A

Une succession de phrases sera projetée sur l'écran. Donnez à la première phrase une valeur d'acceptabilité en utilisant n'importe quel chiffre (à l'exception de 0); ensuite, donnez aux phrases suivantes une valeur proportionnelle à celle que vous avez choisie pour la première phrase. Par exemple, si vous avez donné la valeur 6 à la première phrase et vous pensez que la deuxième phrase est deux fois plus acceptable que la première, mettez 12 pour celle-ci. Ou bien, si vous pensez que la deuxième phrase est trois fois moins acceptable que la première, ne mettez que 2. Par contre, si la deuxième phrase ne vous semble qu'un peu moins acceptable que la première, mettez un 5, et ainsi de suite.
Vous pouvez utiliser le même chiffre pour plusieurs phrases différentes si vous estimez qu'elles ont la même valeur d'acceptabilité.

Exemple

Supposez que la première phrase soit la suivante:

(1) Maria ha mangiato gli spaghetti

Celle-ci est une phrase acceptable et vous choisissez, par exemple, la valeur 10. Ensuite, vous voyez la deuxième phrase:

(2) Maria ha mangiati gli spaghetti

Vous pourriez penser que cette phrase est deux fois moins acceptable que la précédente: vous mettez donc 5.
La troisième phrase est:

(3) Maria è gli spaghetti mangiato

Cette phrase pourrait vous sembler encore moins acceptable que (2): donc vous lui donnez la valeur3.

Plus la phrase vous semble acceptable, plus élevée sera le chiffre que vous choisirez pour indiquer sa valeur.

Test B

Vous avez une série de fiches, chacune avec une phrase imprimée. Votre tâche est de ranger les phrases dans des piles différentes selon leur dégré d'acceptabilité. Vous pouvez former autant de piles que vous voulez, même si chaque pile ne contient qu'une seule phrase. Quand vous aurez distribué toutes les phrases dans des piles, étalez les piles sur votre bureau, en plaçant les moins acceptables à gauche et les plus acceptables à droite. Ensuite, utilisez le papier millimétré pour faire une liste des phrases que vous avez mises dans chaque pile et pour indiquer l'ordre dans lequel vous avez arrangé les piles. Faites ceci en marquant les chiffres qui correspondent aux phrases dans chaque pile dans des colonnes séparées, et dans le même ordre que sur votre bureau ("moins acceptable" à gauche; "plus acceptable" à droite).

Exemples

Supposez que sept des fiches correspondent aux phrases suivantes:

(1) Maria ha mangiato gli spaghetti
(2) Maria è gli spaghetti mangiato
(3) E' mangiare gli spaghetti che è piaciuto a Maria
(4) Di spaghetti, Maria ha mangiati molti
(5) Di spaghetti, Maria ne ha mangiati molti
(6) E' mangiati gli spaghetti che sono piaciuti a Maria
(7) Maria ha mangiati gli spaghetti

Les phrases (1), (3) et (5) sont tout à fait acceptables: elles appartiennent donc à la même pile. Les phrases (2) et (6) sont complètement inacceptables, et peuvent donc être placées dans la même pile. Les phrases (4) et (7) sont moins acceptables que celles de la première pile, mais plus acceptables que celles de la deuxième pile: elles forment une pile séparée (pourtant, si (4) vous semble moins acceptable que (7), mettez ces phrases dans deux piles différentes!).

Vos piles seront arrangées de la manière suivante:

2	4	1
6	7	3
		5

Enfin, vous marquerez les numéros des phrases dans le même ordre, comme dans l'example suivant:

MAIN STUDY

Sentences / Lexicalization 1
(sentences correspond to structure types in Fig. 6.1)

1. Mia sorella è nuotata nella piscina
2. Mia figlia ha nuotato tutti i giorni quest'estate
3. Di campioni olimpionici, ne sono nuotati molti in questa piscina
4. Di campioni olimpionici ne hanno nuotato molti in questa piscina
5. I ragazzi sono giocati tutto il pomeriggio
6. Dopo cena gli ospiti hanno giocato a carte
7. Di ragazzi ne sono giocati molti in questo giardino
8. Di ragazzi, ne hanno giocato molti in questo giardino
9. Paola è corsa più velocemente di Maria
10. Paola ha corso più velocemente di tutti
11. Di atleti, ne sono corsi molti sotto effetto di droghe
12. Di atleti, ne hanno corso molti sotto effetto di droghe

13. Maria è venuta alla festa con il suo amico
14. Carla ha venuto a teatro con noi
15. Di stranieri in Italia, ne sono venuti molti l'anno scorso
16. Di stranieri in Italia, ne hanno venuto molti l'anno scorso
17. Carla è restata da me fino a tardi
18. Maria ha restato a Firenze tutto l'inverno
19. Di stranieri in Italia, ne sono restati molti
20. Di stranieri in Italia, ne hanno restato molti
21. I dinosauri sono esistiti milioni di anni fa
22. Gli unicorni non hanno mai esistito
23. Di dinosauri, ne sono esistiti migliaia di specie
24. Di dinosauri, ne hanno esistito migliaia di specie
25. Le tasse sono aumentate del 20%
26. I prezzi hanno aumentato del 10%
27. Di prezzi, ne sono aumentati molti dopo l'estate
28. Di prezzi, ne hanno aumentato molti dopo l'estate
29. Paola è corsa in farmacia
30. Maria ha corso a casa
31. Di gatti dal Colosseo, ne sono corsi via molti
32. Di gatti dal Colosseo, ne hanno corso via molti

33. Maria ha potuto andare alla tua festa
34. E' andare alla tua festa che Maria non ha potuto
35. Alla tua festa, Maria non ci ha potuto andare
36. Alla tua festa, Maria non ha potuto andarci
37. Mia figlia non è potuta andare a scuola stamattina
38. E' andare a scuola che mia figlia non è potuta
39. A scuola, mia figlia non ci è potuta andare
40. A scuola, mia figlia non è potuta andarci
41. La famosa attrice ha voluto tornare in albergo da sola
42. E' tornare in albergo che la famosa attrice ha voluto
43. In albergo, la famosa attrice ci ha voluto tornare da sola
44. In albergo, la famosa attrice ha voluto tornarci da sola
45. Dopo l'incidente, Paola non è voluta tornare a casa
46. E' tornare a casa dopo l'incidente che Paola non è voluta
47. A casa dopo l'incidente, Paola non ci è voluta tornare
48. A casa dopo l'incidente, Paola non è voluta tornarci.

Lexicalization 2
(sentences correspond to structure types in Fig. 6.1)

1. Michela è viaggiata molto in America
2. Mia zia ha viaggiato molto da giovane
3. Di giovani, ne sono viaggiati molti in treno
4. Di giovani, ne hanno viaggiato molti in treno
5. Carla è dormita fino a mezzogiorno
6. Franca ha dormito tutto il pomeriggio
7. Di professori, ne sono dormiti molti durante la conferenza
8. Di professori, ne hanno dormito molti durante la conferenza
9. La bambina è saltata sul letto per un'ora
10. Sara Simeoni ha saltato alle Olimpiadi di Mosca
11. Di atleti italiani, non ne sono saltati molti alle scorse Olimpiadi
12. Di atleti italiani, non ne hanno saltato molti alle scorse Olimpiadi

13. Francesca è arrivata in ritardo all'appuntamento
14. Maria ha arrivato con tre ore di ritardo
15. Di treni in ritardo, ne sono arrivati molti
16. Di treni in ritardo, ne hanno arrivato molti
17. I miei nonni sono sopravvissuti al terremoto del 1908
18. Molti argentini hanno sopravvissuto alle persecuzioni politiche
19. Di prigionieri, ne sono sopravvissuti molti
20. Di prigionieri, ne hanno sopravvissuto molti
21. Questa pipa è appartenuta a mio padre
22. Questa casa ha appartenuto ai miei nonni
23. A mio padre, di pipe, gliene sono appartenute molte
24. A mio padre, di pipe, gliene hanno appartenuto molte
25. La nave è affondata rapidamente
26. Le navi hanno affondato a 3 Km dal porto
27. Di transatlantici, non ne sono mai affondati molti
28. Di transatlantici, non ne hanno mai affondato molti
29. I bambini sono saltati giù dal letto
30. I ladri hanno saltato giù dalla finestra
31. Di ragazzi, ne sono saltati giù molti da questo muro
32. Di ragazzi, ne hanno saltato giù molti da questo muro

33. Francesca ha dovuto rimanere a casa tutta la sera
34. E' rimanere a casa tutta la sera che Francesca ha dovuto
35. A casa, Francesca ci ha dovuto rimanere tutta la sera
36. A casa, Francesca ha dovuto rimanerci tutta la sera
37. Giovanna è dovuta rimanere a Roma tutta l'estate
38. E' rimanere a Roma tutta l'estate che Giovanna è dovuta
39. A Roma, Giovanna ci è dovuta rimanere tutta l'estate
40. A Roma, Giovanna è dovuta rimanerci tutta l'estate
41. Maria non ha saputo arrivare a casa tua da sola
42. E' arrivare a casa tua che Maria non ha saputo
43. Dove si trova la tua casa? Maria non ci ha saputo arrivare da sola
44. Dove si trova la tua casa? Maria non ha saputo arrivarci da sola
45. Paola non è saputa arrivare al cinema
46. E' arrivare al cinema che Paola non è saputa
47. Dove si trova il cinema? Paola non ci è saputa arrivare
48. Dove si trova il cinema? Paola non è saputa arrivarci

Lexicalization 3
(sentences correspond to structure types in Fig. 6.1)

1. Siamo camminati nel parco tutto il pomeriggio
2. Abbiamo camminato in campagna per tre ore
3. Di attori famosi, ne sono camminati molti in via Veneto
4. Di attori famosi, ne hanno camminato molti nel centro di Roma
5. Paola è lavorata in un pub di Londra quest'estate
6. Carla ha lavorato in un ristorante l'anno scorso
7. Di turisti stranieri, ne sono lavorati molti in questo pub
8. Di ragazzi inglesi, ne hanno lavorato molti in questo ristorante
9. Gli elicotteri sono volati a lungo sulla città
10. Gli uccelli hanno volato da un albero all'altro
11. Di aerei, ne sono volati molti a bassa quota sulla città
12. Di uccelli, ne hanno volato molti davanti alla finestra

13. Giovanna è ritornata al suo paese dopo vent'anni
14. Maria ha ritornato ieri dalle vacanze
15. Di immigranti, ne sono ritornati molti in Italia nei primi anni '60
16. Di turisti tedeschi, ne hanno ritornato molti a Capri l'estate scorsa
17. I miei parenti sono rimasti da me fino a tardi
18. I miei amici hanno rimasto al bar fino alle tre del mattino
19. Di negozi, ne sono rimasti molti aperti durante il periodo natalizio
20. Di ristoranti, ne sono rimasti pochi aperti il giorno di Pasqua
21. Quella storia mi è sembrata incredibile
22. L'attrice in quel film mi ha sembrato molto brava
23. Di attrici in quel film, me ne sono sembrate brave solo due
24. delle sue storie, me ne hanno sembrato incredibili molte
25. La casa dei miei genitori non è cambiata dalla mia infanzia
26. La mia vita ha completamente cambiato
27. Di città, ne sono cambiate molte dall'ultima guerra
28. Di programmi televisivi, ne hanno cambiato molti dall'anno scorso
29. I canarini sono volati via dalla gabbia
30. I passerotti hanno volato via dal nido
31. Di uccellini, ne sono volati via dal nido tre
32. Di rondini, ne hanno volato via molte alla fine dell'estate

33. Gianna ha continuato ad andare a lezione di giapponese
34. E' ad andare a lezione di giapponese che Gianna ha continuato
35. A lezione di giapponese, Gianna ci ha continuato ad andare
36. A lezione di giapponese, Gianna ci è continuata ad andare
37. L'infermiera è continuata a venire a casa mia tutti i giorni
38. E' a venire a casa mia tutti i giorni che l'infermiera è continuata
39. A casa mia, l'infermiera ci è continuata a venire tutti i giorni
40. A casa mia, l'infermiera è continuata a venirci tutti i giorni
41. Gli impiegati hanno cominciato ad arrivare in ufficio in ritardo
42. E' ad arrivare in ritardo che gli impiegati hanno cominciato
43. In ufficio, gli impiegati ci hanno cominciato ad arrivare in ritardo
44. In ufficio, gli impiegati hanno cominciato ad arrivarci in ritardo
45. Gli invitati sono cominciati ad arrivare alla festa un'ora fa
46. E' ad arrivare alla festa che gli invitati sono cominciati
47. Alla festa, gli invitati ci sono cominciati ad arrivare un'ora fa
48. Alla festa, gli invitati sono cominciati ad arrivarci un'ora fa

Lexicalization 4
(sentences correspond to structure types in Fig. 6.1)

1. Paola è passeggiata nel giardino con i bambini
2. Carla ha passeggiato nel parco fino all'ora di pranzo
3. Di coppie di innamorati, ne sono passeggiate tante in questo giardino
4. Di mamme con bambini, ne hanno passeggiato molte in questo parco
5. I candidati sono parlati alla televisione per un'ora
6. Gli studenti hanno parlato su Dante e Boccaccio alla conferenza
7. Di studenti, ne sono parlati quattro alla conferenza
8. Di candidati, ne hanno parlato tre alla televisione
9. La pallina è rotolata velocemente
10. La moneta ha rotolato lungo il corridoio
11. Di monete, ne sono rotolate tre
12. Di palline, ne hanno rotolato due dietro al divano

13. Le foglie sono cadute dagli alberi
14. La bambina ha caduto dal seggiolone
15. Di foglie, ne sono cadute molte all'inizio dell'autunno
16. Di alberi, ne sono caduti molti durante l'uragano
17. La tempesta è durata tutta la notte
18. La prima guerra mondiale ha durato tre anni
19. Di governi in Italia, ne sono durati pochi per più di un anno
20. Di mostre al Louvre, ne hanno durato molte per più di un anno
21. La commedia è piaciuta molto al pubblico
22. I film giapponesi hanno piaciuto agli spettatori italiani
23. Di spettacoli teatrali al Festival, ne sono piaciuti molti ai giovani
24. Di dolci siciliani, ne hanno piaciuto molti al bambino
25. I risultati scolastici di Paola sono migliorati notevolmente
26. Le condizioni di vita nel Sud hanno nettamente migliorato
27. Di studenti, ne sono migliorati molti durante l'anno
28. Di primati mondiali alle Olimpiadi, ne hanno migliorato alcuni
29. La pallina da golf è rotolata giù nella buca
30. La moneta ha rotolato giù dalla tasca
31. Di monete, ne sono rotolate due giù dalla tasca
32. Di palline, ne hanno rotolato tre giù nella buca

33. La gente ha continuato ad andare al concerto del lunedì
34. E' ad andare al concerto del lunedì che la gente ha continuato
35. Al concerto del lunedì, la gente ci ha continuato ad andare
36. Al concerto del lunedì, la gente ha continuato ad andarci
37. I turisti tedeschi sono continuati ad andare in vacanza in Italia
38. E' ad andare in vacanza in Italia che i turisti tedeschi sono continuati
39. In vacanza in Italia, i turisti tedeschi ci sono continuati ad andare
40. In vacanza in Italia, i turisti tedeschi sono continuati ad andarci
41. La gente ha cominciato a ritornare a casa dopo mezzanotte
42. E' a ritornare a casa che la gente ha cominciato
43. A casa, la gente ci ha cominciato a ritornare dopo mezzanotte
44. A casa, la gente ha cominciato a ritornarci dopo mezzanotte
45. L'acqua è cominciata a penetrare in casa dal tetto
46. E' a penetrare in casa dal tetto che l'acqua è cominciata
47. In casa, l'acqua ci è cominciata a penetrare dal tetto
48. In casa, l'acqua è cominciata a penetrarci dal tetto

APPENDIX B

ANOVA TABLES

LIST OF ABBREVIATIONS USED IN TABLES:
(NB: abbreviations refer to the names of the main effects; variables in interactions are defined by their first letter. Example: AP = auxiliary x proficiency level)

AUX = auxiliary
NLG = native language
PLEVEL = proficiency level
RC = raising/control
SCAT = semantic or syntactic category
WO = word order

B.1 PILOT STUDIES

B.1.1 PILOT 1

TABLE B.1: PILOT 1. Acceptability judgments of French learners of Italian on Italian intransitive verbs and of Italian learners of French on French intransitive verbs: auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	20596.29285	1	20596.29285	1921.56	0.0000
NLG	0.52174	1	0.52174	0.05	0.8268
PLEVEL	86.28809	2	43.14405	4.03	0.0276
NP	3.39299	2	1.69650	0.16	0.8543
1 ERROR	342.99315	32	10.71854		
SCAT	121.78967	9	13.53219	4.30	0.0000
SN	54.86405	9	6.09601	1.94	0.0468
SP	70.91436	18	3.93969	1.25	0.2195
SNP	70.94679	18	3.94149	1.25	0.2190
2 ERROR	906.49171	288	3.14754		
AUX	1596.77475	1	1596.77475	112.42	0.0000
AN	11.07674	1	11.07674	0.78	0.3838
AP	557.25705	2	278.62853	19.62	0.0000
ANP	50.82817	2	25.41408	1.79	0.1834
3 ERROR	454.53594	32	14.20425		
SA	1392.66526	9	154.74058	27.06	0.0000
SAN	266.09307	9	29.56590	5.17	0.0000
SAP	192.11569	18	10.67309	1.87	0.0184
SANP	196.99817	18	10.94434	1.91	0.0148
4 ERROR	1646.91475	288	5.71845		

B.1.2 PILOT 2

TABLE B.2. PILOT 2. Magnitude Estimation. Transitive and intransitive verbs, auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	31.59970	1	31.59970	63.73	0.0000
PLEVEL	8.97859	3	2.99286	6.04	0.0023
1 ERROR	15.37186	31	0.49587		
SCAT	18.61396	11	1.69218	12.14	0.0000
SP	9.54261	33	0.28917	2.07	0.0007
2 ERROR	47.54917	341	0.13944		

TABLE B.3. PILOT 2. Card-Sorting. Transitive and intransitive verbs, auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	557.10145	1	557.10145	110.37	0.0000
PLEVEL	0.57971	1	0.57971	0.11	0.7381
1 ERROR	106.00000	21	5.04762		
SCAT	210.92947	11	19.17541	16.37	0.0000
SP	13.19034	11	1.19912	1.02	0.4261
2 ERROR	270.57778	231	1.17133		

TABLE B.4. PILOT 2. Ranking. Transitive and intransitive verbs, auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	1971.86191	1	1971.86191	46.21	0.0000
PLEVEL	359.69074	3	119.89691	2.81	0.0557
1 ERROR	1322.77593	31	42.67019		
SCAT	859.04752	11	78.09523	11.25	0.0000
SP	648.17593	33	19.64169	2.83	0.0000
2 ERROR	2368.15741	341	6.94474		

B.1.3 PILOT 3

TABLE B.5. PILOT 3. Magnitude Estimation. Imperfective constructions, auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	12.78425	1	12.78425	77.26	0.0000
PLEVEL	1.70927	3	0.56976	3.44	0.0282
1 ERROR	5.29504	32	0.16547		
SCAT	7.13471	9	0.79275	9.48	0.0000
SP	5.11826	27	0.18957	2.27	0.0005
2 ERROR	24.07098	288	0.08358		

TABLE B.6. PILOT 3. Card-Sorting. Imperfective constructions, auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	1310.72000	1	1310.72000	56.08	0.0000
PLEVEL	113.28889	3	37.76296	1.62	0.2051
1 ERROR	747.86667	32	23.37083		
SCAT	278.96889	9	30.99654	4.12	0.0001
SP	586.60000	27	21.72593	2.88	0.0000
2 ERROR	2168.91111	288	7.53094		

TABLE B.7. PILOT 3. Ranking. Imperfective constructions, auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	2415.12500	1	2415.12500	51.25	0.0000
PLEVEL	398.38056	3	132.79352	2.82	0.0547
1 ERROR	1507.98333	32	47.12448		
SCAT	511.02500	9	56.78056	3.30	0.0008
SP	707.25833	27	26.19475	1.52	0.0508
2 ERROR	4957.51667	288	17.21360		

TABLE B.8. PILOT 3. Magnitude Estimation. Perfective constructions, auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	13.33672	1	13.33672	52.67	0.0000
PLEVEL	3.69082	3	1.23027	4.86	0.0068
1 ERROR	8.10238	32	0.25320		
SCAT	1.60862	4	0.40216	4.66	0.0015
SP	1.76863	12	0.14739	1.71	0.0726
2 ERROR	11.05791	128	0.08639		

TABLE B.9. PILOT 3. Card-Sorting. Perfective constructions, auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	932.28444	1	932.28444	30.48	0.0000
PLEVEL	234.75556	3	78.25185	2.56	0.0724
1 ERROR	978.82222	32	30.58819		
SCAT	295.90444	4	73.97611	9.39	0.0000
SP	207.07778	12	17.25648	2.19	0.0157
2 ERROR	1008.28889	128	7.87726		

TABLE B.10. PILOT 3. Ranking. Perfective constructions, auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	1705.69000	1	1705.69000	28.03	0.0000
PLEVEL	756.69444	3	252.23148	4.15	0.0137
1 ERROR	1947.23333	32	60.85104		
SCAT	383.81556	4	95.95389	5.59	0.0004
SP	195.16667	12	16.26389	0.95	0.5025
2 ERROR	2197.37778	128	17.16701		

B.2 MAIN STUDY

B.2.1 MAGNITUDE ESTIMATION

B.2.1.1 BY SUBJECTS (TABLES 1-12)
B.2.1.2 BY MATERIALS (TABLES 13-24)

B.2.1.1 BY SUBJECTS

TABLE B.11: Magnitude Estimation. Unergative verbs, all levels, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	5695.26300	1	5695.26300	3839.81	0.0000
PLEVEL	75.55572	4	18.88893	12.74	0.0000
1 ERROR	229.89801	155	1.48321		
SCAT	7.35789	2	3.67894	18.69	0.0000
SP	8.37168	8	1.04646	5.32	0.0000
2 ERROR	61.02280	310	0.19685		
WO	12.68638	1	12.68638	44.28	0.0000
WP	12.16349	4	3.04087	10.61	0.0000
3 ERROR	44.40662	155	0.28649		
SW	0.09647	2	0.04824	0.32	0.7286
SWP	0.98433	8	0.12304	0.81	0.5954
4 ERROR	47.17049	310	0.15216		
AUX	77.44616	1	77.44616	165.78	0.0000
AP	11.99663	4	2.99916	6.42	0.0001
5 ERROR	72.40931	155	0.46716		
SA	42.34052	2	21.17026	70.89	0.0000
SAP	14.87669	8	1.85959	6.23	0.0000
6 ERROR	92.57326	310	0.29862		
WA	24.66890	1	24.66890	81.63	0.0000
WAP	18.14791	4	4.53698	15.01	0.0000
7 ERROR	46.84132	155	0.30220		
SWA	2.69006	2	1.34503	9.08	0.0001
SWAP	2.77904	8	0.34738	2.34	0.0185
8 ERROR	45.92238	310	0.14814		

TABLE B.12: Magnitude Estimation. Unergative verbs, all levels, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	154.89232	1	154.89232	165.78	0.0000
PLEVEL	23.99326	4	5.99831	6.42	0.0001
1 ERROR	144.81861	155	0.93431		
SCAT	84.68103	2	42.34052	70.89	0.0000
SP	29.75339	8	3.71917	6.23	0.0000
2 ERROR	185.14653	310	0.59725		
WO	49.33781	1	49.33781	81.63	0.0000
WP	36.29583	4	9.07396	15.01	0.0000
3 ERROR	93.68265	155	0.60440		
SW	5.38012	2	2.69006	9.08	0.0001
SWP	5.55809	8	0.69476	2.34	0.0185
4 ERROR	91.84476	310	0.29627		

TABLE B.13: Magnitude Estimation. Unaccusative verbs, all levels, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	1795.99578	1	1795.99578	4982.33	0.0000
PLEVEL	5.24859	4	1.31215	3.64	0.0073
1 ERROR	55.87337	155	0.36047		
SCAT	5.68553	4	1.42138	29.26	0.0000
SP	1.83142	16	0.11446	2.36	0.0021
2 ERROR	30.11355	620	0.04857		
WO	4.42453	1	4.42453	67.34	0.0000
WP	1.14425	4	0.28606	4.35	0.0023
3 ERROR	10.18424	155	0.06570		
SW	0.08058	4	0.02015	0.54	0.7078
SWP	0.31112	16	0.01944	0.52	0.9378
4 ERROR	23.21191	620	0.03744		
AUX	54.07113	1	54.07113	366.03	0.0000
AP	17.07991	4	4.26998	28.91	0.0000
5 ERROR	22.89686	155	0.14772		
SA	16.31928	4	4.07982	73.17	0.0000
SAP	0.58579	16	0.03661	0.66	0.8372
6 ERROR	34.56821	620	0.05576		
WA	0.22855	1	0.22855	4.66	0.0325
WAP	0.74071	4	0.18518	3.77	0.0059
7 ERROR	7.61015	155	0.04910		
SWA	0.65194	4	0.16299	3.79	0.0047
SWAP	0.56433	16	0.03527	2.82	0.0063
8 ERROR	26.62963	620	0.04295		

TABLE B.14: Magnitude Estimation / Unaccusative verbs, all levels, auxiliary preferences

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	108.14225	1	108.14225	366.03	0.0000
PLEVEL	34.15981	4	8.53.995	28.91	0.0000
1 ERROR	45.79372	155	0.29544		
SCAT	32.63857	4	8.15964	73.17	0.0000
SP	1.17157	16	0.07322	0.66	0.8372
2 ERROR	69.13642	620	0.11151		
WO	0.45711	1	0.45711	4.66	0.0325
WP	1.48141	4	0.37035	3.77	0.0059
3 ERROR	15.22029	155	0.09820		
SW	1.30388	4	0.32597	3.79	0.0047
SWP	1.12866	16	0.07054	2.82	0.0063
4 ERROR	106.51852	620	0.08454		

TABLE B.15: Magnitude Estimation. Restructuring verbs, all levels, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	1441.33423	1	1441.33423	3997.59	0.0000
PLEVEL	2.15966	4	0.53991	1.50	0.2056
1 ERROR	55.88532	155	0.36055		
RC	0.00888	1	0.00888	0.23	0.6329
RP	0.31906	4	0.07977	2.06	0.0891
2 ERROR	6.01064	155	0.03878		
SCAT	1.43038	1	1.43038	13.96	0.0003
SP	0.43532	4	0.10883	1.06	0.3774
3 ERROR	15.88401	155	0.10248		
RS	0.60891	1	0.60891	11.06	0.0011
RSP	0.08473	4	0.02118	0.38	0.8193
4 ERROR	8.53244	155	0.05505		
WO	33.05459	3	11.01820	150.81	0.0000
WP	3.43434	12	0.28619	3.92	0.0000
5 ERROR	33.97266	465	0.07306		
RW	0.09415	3	0.03138	1.09	0.3539
RWP	0.40633	12	0.03386	1.17	0.2993
6 ERROR	13.41699	465	0.02885		
SW	1.17516	3	0.39172	8.52	0.0000
SWP	6.77245	12	0.56437	12.27	0.0000
7 ERROR	21.38914	465	0.04600		
RSW	0.31237	3	0.10412	2.59	0.0525
RSWP	0.26305	12	0.02192	0.54	0.8852
8 ERROR	18.71227	465	0.04024		

TABLE B.16: Magnitude Estimation. Restructuring verbs, all levels, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	4.31999	1	4.31999	31.10	0.0000
PLEVEL	6.39921	4	1.59980	11.52	0.0000
1 ERROR	21.53315	155	0.13892		
RC	0.24441	1	0.24441	2.96	0.0874
RP	0.07675	4	0.01919	0.23	0.9199
2 ERROR	12.80101	155	0.08259		
WO	0.89108	3	0.29703	2.61	0.0513
WP	8.01632	12	0.66803	5.86	0.0000
3 ERROR	53.01314	465	0.11401		
RW	1.59814	3	0.53271	5.94	0.0006
RWP	0.61881	12	0.05157	0.58	0.8625
4 ERROR	41.68840	465	0.08965		

TABLE B.17: Magnitude Estimation. Unergative verbs, native vs near-native, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	2179.96233	1	2179.96233	1009.89	0.0000
NLG	20.19657	2	10.09829	4.68	0.0121
1 ERROR	166.21272	77	2.15861		
SCAT	15.44668	2	7.72334	28.30	0.0000
SN	3.14091	4	0.78523	2.88	0.0247
2 ERROR	42.03204	154	0.27294		
WO	17.79162	1	17.79162	46.12	0.0000
WN	5.74968	2	2.87484	7.45	0.0011
3 ERROR	29.70277	77	0.38575		
SW	1.45157	2	0.72578	3.69	0.0271
SWN	1.05797	4	0.26449	1.35	0.2554
4 ERROR	30.25897	154	0.19649		
AUX	86.14327	1	86.14327	157.05	0.0000
AN	0.33079	2	0.16539	0.30	0.7406
5 ERROR	42.23541	77	0.54851		
SA	49.64857	2	24.82429	61.47	0.0000
SAN	0.97010	4	0.24253	0.60	0.6628
6 ERROR	62.18934	154	0.40383		
WA	38.47342	1	38.47342	91.60	0.0000
WAN	7.93892	2	3.96946	9.45	0.0002
7 ERROR	32.34269	77	0.42003		
SWA	3.37415	2	1.68708	9.44	0.0001
SWAN	0.38037	4	0.09509	0.53	0.7124
8 ERROR	27.52614	154	0.17874		

TABLE B.18: Magnitude Estimation. Unergative verbs, native vs near-native, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	172.28655	1	172.28655	157.05	0.0000
NLG	0.66158	2	0.33079	0.30	0.7406
1 ERROR	84.47082	77	1.09702		
SCAT	99.29714	2	49.64857	61.47	0.0000
SN	1.94021	4	0.48505	0.60	0.6628
2 ERROR	124.37869	154	0.80765		
WO	76.94684	1	76.94684	91.60	0.0000
WN	15.87784	2	7.93892	9.45	0.0002
3 ERROR	64.68538	77	0.84007		
SW	6.74830	2	3.37415	9.44	0.0001
SWN	0.76074	4	0.19019	0.53	0.7124
4 ERROR	55.05229	154	0.35748		

TABLE B.19: Magnitude Estimation. Unaccusative verbs, native vs near-native, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	787.41200	1	787.41200	1825.88	0.0000
NLG	1.19500	2	0.59750	1.39	0.2564
1 ERROR	33.20632	77	0.43125		
SCAT	7.17780	4	1.79445	32.85	0.0000
SN	0.33042	8	0.04130	0.76	0.6419
2 ERROR	16.82565	308	0.05463		
WO	2.61366	1	2.61366	33.18	0.0000
WN	1.49412	2	0.74706	9.49	0.0002
3 ERROR	6.06462	77	0.07876		
SW	0.38671	4	0.09668	2.17	0.0721
SWN	0.55893	8	0.06987	1.57	0.1332
4 ERROR	13.71133	308	0.04452		
AUX	58.69481	1	58.69481	337.06	0.0000
AN	2.33251	2	1.16625	6.70	0.0021
5 ERROR	13.40842	77	0.17414		
SA	12.99733	4	3.24933	61.16	0.0000
SAN	0.57708	8	0.07213	1.36	0.2146
6 ERROR	16.36338	308	0.05313		
WA	0.00537	1	0.00537	0.08	0.7757
WAN	0.11829	2	0.05915	0.90	0.4107
7 ERROR	5.05914	77	0.06570		
SWA	0.40408	4	0.10102	1.95	0.1021
SWAN	0.31958	8	0.03995	4.77	0.0372
8 ERROR	15.95543	308	0.05180		

TABLE B.20: Magnitude Estimation. Unaccusative verbs, native vs near-native, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	117.38961	1	117.38961	337.06	0.0000
NLG	4.66502	2	2.33251	6.70	0.0021
1 ERROR	26.81684	77	0.34827		
SCAT	25.99465	4	6.49866	61.16	0.0000
SN	1.15415	8	0.14427	1.36	0.2146
2 ERROR	65.45353	308	0.10423		
WO	0.01074	1	0.01074	0.08	0.7757
WN	0.23658	2	0.11829	0.90	0.4107
3 ERROR	10.11828	77	0.13141		
SW	10.80817	4	0.20204	1.95	0.1021
SWN	0.63915	8	0.07989	4.77	0.0372
4 ERROR	31.91086	308	0.10361		

TABLE B.21: Magnitude Estimation. Restructuring verbs, native vs near-native, all variables (BYSUBJECT)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	687.24952	1	687.24952	1609.10	0.0000
NLG	0.87195	2	0.43597	1.02	0.3651
1 ERROR	32.88692	77	0.42710		
RC	0.07423	1	0.07423	1.45	0.2330
RN	0.16426	2	0.08213	1.60	0.2088
2 ERROR	3.95570	77	0.05137		
SCAT	1.37219	1	1.37219	28.20	0.0000
SN	0.81306	2	0.40653	8.36	0.0005
3 ERROR	3.74643	77	0.04865		
RS	0.26744	1	0.26744	4.25	0.0426
RSN	0.05341	2	0.02671	0.42	0.6556
4 ERROR	4.84498	77	0.06292		
WO	16.93516	3	5.64505	67.42	0.0000
WN	2.79462	6	0.46577	5.56	0.0000
5 ERROR	19.34132	231	0.08373		
RW	0.35602	3	0.11867	3.39	0.0189
RWN	1.18428	6	0.19738	5.63	0.0000
6 ERROR	8.09281	231	0.03503		
SW	7.08901	3	2.36300	36.69	0.0000
SWN	4.52420	6	0.75403	11.71	0.0000
7 ERROR	14.87656	231	0.06440		
RSW	0.15222	3	0.05074	1.06	0.3653
RSWN	0.81354	6	0.13559	2.84	0.0109
8 ERROR	11.02016	231	0.04771		

TABLE B.22: Magnitude Estimation. Restructuring verbs, native vs near-native, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	16.61957	1	16.61957	81.61	0.0000
NLG	6.07423	2	3.03712	14.91	0.0000
1 ERROR	15.68040	77	0.20364		
RC	0.05588	1	0.05588	0.50	0.4807
RN	0.08041	2	0.04021	0.36	0.6980
2 ERROR	8.57016	77	0.11130		
WO	0.30284	3	0.10095	1.08	0.3577
WN	4.60029	6	0.76672	8.21	0.0000
3 ERROR	21.56559	231	0.09336		
RW	0.78345	3	0.26115	2.60	0.0526
RWN	1.65350	6	0.27558	2.75	0.0134
4 ERROR	23.16012	231	0.10026		

B.2.1.2 BY MATERIALS

TABLE B.23: Magnitude Estimation. Unergative verbs, all levels, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	727.72781	1	727.72781	10584.16	0.0000
SCAT	0.94018	2	0.47009	6.84	0.0062
WO	1.62102	1	1.62102	23.58	0.0001
SW	0.01233	2	0.00616	0.09	0.9147
1 ERROR	1.23761	18	0.06876		
PLEVEL	8.68290	4	2.17073	29.82	0.0000
PS	0.97049	8	0.12131	1.67	0.1215
PW	1.40340	4	0.35085	4.82	0.0017
PSW	0.11555	8	0.01444	0.20	0.9902
2 ERROR	5.24059	72	0.07279		
AUX	9.89587	1	9.89587	86.46	0.0000
AS	5.41015	2	2.70508	23.64	0.0000
AW	3.15212	1	3.15212	27.54	0.0001
ASW	0.34372	2	0.17186	1.50	0.2494
3 ERROR	2.06010	18	0.11445		
PA	1.60624	4	0.40156	10.82	0.0000
PAS	1.91052	8	0.23882	6.44	0.0000
PAW	2.05650	4	0.51412	13.85	0.0000
PASW	0.34875	8	0.04359	1.17	0.3263
4 ERROR	2.67204	72	0.03711		

TABLE B.24: Magnitude Estimation. Unergative verbs, all levels, auxiliary differences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	19.79175	1	19.79175	86.46	0.0000
SCAT	10.82030	2	5.41015	23.64	0.0000
WO	6.30424	1	6.30424	27.54	0.0001
SW	0.68744	2	0.34372	1.50	0.2494
1 ERROR	4.12020	18	0.22890		
PLEVEL	3.21247	4	0.80312	10.82	0.0000
PS	3.82104	8	0.47763	6.44	0.0000
PW	4.11300	4	1.02825	13.85	0.0000
PSW	0.69750	8	0.08719	1.17	0.3263
2 ERROR	5.34408	72	0.07422		

TABLE B.25: Magnitude Estimation. Unaccusative verbs, all levels, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	1221.99082	1	1221.99082	3961.55	0.0000
SCAT	3.71212	4	0.92803	3.01	0.0336
WO	3.24070	1	3.24070	10.51	0.0029
SW	0.07179	4	0.01795	0.06	0.9934
1 ERROR	9.25387	30	0.30846		
PLEVEL	3.18295	4	0.79574	9.31	0.0000
PS	1.14514	16	0.07157	0.84	0.6410
PW	0.67395	4	0.16849	1.97	0.1031
PSW	0.22322	16	0.01395	0.16	0.9999
2 ERROR	10.25364	120	0.08545		
AUX	36.30342	1	36.30342	82.13	0.0000
AS	11.25868	4	2.81467	6.37	0.0008
AW	0.17148	1	0.17148	0.39	0.5381
ASW	0.48073	4	0.12018	0.27	0.8937
3 ERROR	13.25997	30	0.44200		
PA	10.38091	4	2.59523	55.03	0.0000
PAS	0.51494	16	0.03218	0.68	0.8064
PAW	0.45986	4	0.11496	2.44	0.0507
PASW	0.32679	16	0.02042	0.43	0.9708
4 ERROR	5.65879	120	0.04716		

TABLE B.26: Magnitude Estimation. Unaccusative verbs, all levels, auxiliary preferences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	72.60683	1	72.60683	82.13	0.0000
SCAT	22.51736	4	5.62934	6.37	0.0008
WO	0.34297	1	0.34297	0.39	0.5381
SW	0.96145	4	0.24036	0.27	0.8937
1 ERROR	26.51995	30	0.88400		
PLEVEL	20.76181	4	5.19045	55.03	0.0000
PS	1.02989	16	0.06437	0.68	0.8064
PW	0.91972	4	0.22993	2.44	0.0507
PSW	0.65358	16	0.04085	0.43	0.9708
2 ERROR	11.31758	120	0.09431		

TABLE B.27: Magnitude Estimation. Restructuring verbs, all levels, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	985.42103	1	985.42103	9369.13	0.0000
RC	0.02140	1	0.02140	0.20	0.6560
WO	23.11698	3	7.70566	73.26	0.0000
RW	0.07024	3	0.02341	0.22	0.8797
1 ERROR	2.52426	24	0.10518		
PLEVEL	1.33998	4	0.33500	5.52	0.0005
PR	0.17353	4	0.04338	0.72	0.5836
PW	2.18487	12	0.18207	3.00	0.0013
PRW	0.23462	12	0.01955	0.32	0.9837
2 ERROR	5.82341	96	0.06066		
SCAT	1.08485	1	1.08485	10.99	0.0029
SR	0.43502	1	0.43502	4.41	0.0465
SW	1.07833	3	0.35944	3.64	0.0270
SRW	0.16836	3	0.05612	0.57	0.6413
3 ERROR	2.37008	24	0.09875		
PS	0.28025	4	0.07006	1.34	0.2591
PSR	0.06460	4	0.01615	0.31	0.8707
PSW	3.85575	12	0.32131	6.17	0.0000
PSRW	0.19833	12	0.01653	0.32	0.9848
4 ERROR	5.00258	96	0.05211		

TABLE B.28: Magnitude Estimation. Restructuring verbs, all levels, auxiliary differences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	3.57236	1	3.57236	18.09	0.0003
RC	0.11082	1	0.11082	0.56	0.4611
WO	0.75401	3	0.25134	1.27	0.3062
RW	1.09596	3	0.36532	1.85	0.1652
1 ERROR	4.74016	24	0.19751		
PLEVEL	3.63634	4	0.90908	8.72	0.0000
PR	0.08975	4	0.02244	0.22	0.9294
PW	4.63566	12	0.38631	3.71	0.0001
PRW	0.43612	12	0.03634	0.35	0.9773
2 ERROR	10.00517	96	0.10422		

TABLE B.29: Magnitude Estimation. Unergative verbs, native vs near-native, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	347.17901	1	347.17901	4712.25	0.0000
SCAT	2.46004	2	1.23002	16.70	0.0001
WO	2.83346	1	2.83346	38.46	0.0000
SW	0.23118	2	0.11559	1.57	0.2355
1 ERROR	1.32617	18	0.07368		
NLG	2.79229	2	1.39615	18.17	0.0000
NS	0.46479	4	0.11620	1.51	0.2191
NW	0.78815	2	0.39407	5.13	0.0110
NSW	0.19010	4	0.04752	0.62	0.6521
2 ERROR	2.76559	36	0.07682		
AUX	13.71908	1	13.71908	98.00	0.0000
AS	7.90698	2	3.95349	28.24	0.0000
AW	6.12724	1	6.12724	43.77	0.0000
ASW	0.53736	2	0.26868	1.92	0.1756
3 ERROR	2.51972	18	0.13998		
NA	0.04640	2	0.02320	0.46	0.6339
NAS	0.14366	4	0.03592	0.71	0.5874
NAW	1.16178	2	0.58089	11.56	0.0001
NASW	0.06939	4	0.01735	0.35	0.8456
4 ERROR	1.80903	36	0.05025		

TABLE B.30: Magnitude Estimation. Unergative verbs, native vs near-native, auxiliary preferences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	27.43816	1	27.43816	98.00	0.0000
SCAT	15.81395	2	7.90698	28.24	0.0000
WO	12.25449	1	12.25449	43.77	0.0000
SW	1.07471	2	0.53736	1.92	0.1756
1 ERROR	5.03944	18	0.27997		
NLG	0.09279	2	0.04640	0.46	0.6339
NS	0.28732	4	0.07183	0.71	0.5874
NW	2.32355	2	1.16178	11.56	0.0001
NSW	0.13878	4	0.03470	0.35	0.8456
2 ERROR	3.61806	36	0.10050		

TABLE B.31: Magnitude Estimation. Unaccusative verbs, native vs near-native, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	663.39345	1	663.39345	2604.62	0.0000
SCAT	5.44088	4	1.36022	5.34	0.0023
WO	2.09800	1	2.09800	8.24	0.0075
SW	0.13693	4	0.03423	0.13	0.9684
1 ERROR	7.64095	30	0.25470		
NLG	1.03137	2	0.51569	4.94	0.0104
NS	0.22130	8	0.02766	0.26	0.9748
NW	0.89976	2	0.44988	4.31	0.0179
NSW	0.41655	8	0.05207	0.50	0.8526
2 ERROR	6.26792	60	0.10447		
AUX	51.23814	1	51.23814	227.76	0.0000
AS	10.76996	4	2.69249	11.97	0.0000
AW	0.03761	1	0.03761	0.17	0.6855
ASW	0.58926	4	0.14732	0.65	0.6281
3 ERROR	6.74899	30	0.22497		
NA	1.42806	2	0.71403	9.57	0.0002
NAS	0.48168	8	0.06021	0.81	0.5988
NAW	0.02000	2	0.01000	0.13	0.8748
NASW	0.18982	8	0.02373	0.32	0.9562
4 ERROR	4.47552	60	0.07459		

TABLE B.32: Magnitude Estimation. Unaccusative verbs, native vs near-native, auxiliary preferences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	102.47628	1	102.47628	227.76	0.0000
SCAT	21.53991	4	5.38498	11.97	0.0000
WO	0.07522	1	0.07522	0.17	0.6855
SW	1.17852	4	0.29463	0.65	0.6281
1 ERROR	13.49798	30	0.44993		
NLG	2.85613	2	1.42806	9.57	0.0002
NS	0.96335	8	0.12042	0.81	0.5988
NW	0.04000	2	0.02000	0.13	0.8748
NSW	0.37965	8	0.04746	0.32	0.9562
2 ERROR	8.95104	60	0.14918		

TABLE B.33: Magnitude Estimation. Restructuring verbs, native vs near-native, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	584.73557	1	584.73557	3470.84	0.0000
RC	0.05611	1	0.05611	0.33	0.5693
WO	14.33619	3	4.77873	28.37	0.0000
RW	0.20273	3	0.06758	0.40	0.7535
1 ERROR	4.04331	24	0.16847		
NLG	0.55424	2	0.27712	4.16	0.0215
NR	0.13157	2	0.06578	0.99	0.3798
NW	2.22628	6	0.37105	5.57	0.0002
NRW	0.78456	6	0.13076	1.96	0.0895
2 ERROR	3.19599	48	0.06658		
SCAT	1.19985	1	1.19985	7.59	0.0110
SR	0.31549	1	0.31549	2.00	0.1705
SW	5.58088	3	1.86029	11.77	0.0001
SRW	0.11166	3	0.03722	0.24	0.8707
3 ERROR	3.79330	24	0.15805		
NS	0.71454	2	0.35727	5.73	0.0059
NSR	0.01771	2	0.00885	0.14	0.8680
NSW	3.51968	6	0.58661	9.41	0.0000
NSRW	0.63342	6	0.10557	1.69	0.1433
4 ERROR	2.99386	48	0.06237		

TABLE B.34: Magnitude Estimation. Restructuring verbs, native vs near-native, auxiliary differences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	13.32702	1	13.32702	42.16	0.0000
RC	0.04718	1	0.04718	0.15	0.7027
WO	0.23446	3	0.07815	0.25	0.8625
RW	0.80712	3	0.26904	0.85	0.4797
1 ERROR	7.58661	24	0.31611		
NLG	5.05087	2	2.52543	20.24	0.0000
NR	0.08245	2	0.04123	0.33	0.7202
NW	3.41759	6	0.56960	4.57	0.0010
NRW	1.21980	6	0.20330	1.63	0.1595
2 ERROR	5.98773	48	0.12474		

B.2.2 CARD-SORTING

B.2.2.1 BY SUBJECTS (TABLES 35-46)
B.2.2.2 BY MATERIALS (TABLES 47-58)

B.2.2.1 BY SUBJECTS

TABLE B.35: Card-Sorting. Unergative verbs, all levels, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	21631.46742	1	21631.46742	1553.37	0.0000
PLEVEL	99.89867	4	24.97467	1.79	0.1329
1 ERROR	2158.45081	155	13.92549		
SCAT	9.30829	2	4.65414	3.63	0.0276
SP	71.02477	8	8.87810	6.93	0.0000
2 ERROR	397.34606	310	1.28176		
WO	267.07337	1	267.07337	82.65	0.0000
WP	115.31626	4	28.82907	8.92	0.0000
3 ERROR	500.87905	155	3.23148		
SW	2.32228	2	1.16114	0.78	0.4587
SWP	3.21273	8	0.40159	0.27	0.9752
4 ERROR	460.67477	310	1.48605		
AUX	733.96742	1	733.96742	159.16	0.0000
AP	199.73756	4	49.93439	10.83	0.0000
5 ERROR	714.79109	155	4.61156		
SA	400.96484	2	200.48242	67.35	0.0000
SAP	141.54005	8	17.69251	5.94	0.0000
6 ERROR	922.77662	310	2.97670		
WA	302.96145	1	302.96145	163.57	0.0000
WAP	116.09959	4	29.02490	15.67	0.0000
7 ERROR	287.08738	155	1.85218		
SWA	8.00107	2	4.00053	3.92	0.0209
SWAP	10.70579	8	1.33822	1.31	0.2376
8 ERROR	316.64005	310	1.02142		

TABLE B.36: Card-Sorting. Unergative verbs, all levels, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	1467.93483	1	1467.93483	159.16	0.0000
PLEVEL	399.47512	4	99.86878	10.83	0.0000
1 ERROR	1429.58218	155	9.22311		
SCAT	801.92967	2	400.96484	67.35	0.0000
SP	283.08009	8	35.38501	5.94	0.0000
2 ERROR	1845.55324	310	5.95340		
WO	605.92291	1	605.92291	163.57	0.0000
WP	232.19919	4	58.04980	15.67	0.0000
3 ERROR	574.17477	155	3.70435		
SW	16.00214	2	8.00107	3.92	0.0209
SWP	21.41157	8	2.67645	1.31	0.2376
4 ERROR	633.28009	310	2.04284		

TABLE B.37: Card-Sorting. Unaccusative verbs, all levels, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	38928.26691	1	38928.26691	1698.88	0.0000
PLEVEL	197.56722	4	49.39181	2.16	0.0766
1 ERROR	3551.67153	155	22.91401		
SCAT	62.13002	4	15.53250	10.16	0.0000
SP	91.86854	16	5.74178	3.76	0.0000
2 ERROR	947.93958	620	1.52893		
WO	445.94241	1	445.94241	132.87	0.0000
WP	19.49639	4	4.87410	1.45	0.2195
3 ERROR	520.19861	155	3.35612		
SW	13.94584	4	3.48646	3.05	0.0167
SWP	20.99632	16	1.31227	1.15	0.3071
4 ERROR	709.24931	620	1.14395		
AUX	2273.58068	1	2273.58068	233.35	0.0000
AP	884.11722	4	221.02931	22.69	0.0000
5 ERROR	1510.17153	155	9.74304		
SA	416.22831	4	104.05708	39.44	0.0000
SAP	42.51437	16	2.65715	1.01	0.4470
6 ERROR	1635.88750	620	2.63853		
WA	1.38601	1	1.38601	0.86	0.3549
WAP	35.39583	4	8.84896	5.50	0.0004
7 ERROR	249.50417	155	1.60970		
SWA	34.31346	4	8.57836	7.07	0.0000
SWAP	27.39687	16	1.71230	1.41	0.1295
8 ERROR	751.78750	620	1.21256		

TABLE B.38: Card-Sorting. Unaccusative verbs, all levels, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	4547.16135	1	4547.16135	233.35	0.0000
PLEVEL	1768.23444	4	442.05861	22.69	0.0000
1 ERROR	3020.34306	155	19.48608		
SCAT	832.45663	4	208.11416	39.44	0.0000
SP	85.02875	16	5.31430	1.01	0.4470
2 ERROR	3271.77500	620	5.27706		
AUX	2.77201	1	2.77201	0.86	0.3549
AP	70.79167	4	17.69792	5.50	0.0004
3 ERROR	499.00833	155	3.21941		
SA	68.62692	4	17.15673	7.07	0.0000
SAP	54.79375	16	3.42461	1.41	0.1295
4 ERROR	1503.57500	620	2.42512		

TABLE B.39: Card-Sorting. Restructuring verbs, all levels, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	27625.00228	1	27625.00228	1270.14	0.0000
PLEVEL	567.98225	4	141.99556	6.53	0.0001
1 ERROR	3371.17361	155	21.74951		
RC	0.46515	1	0.46515	0.24	0.6235
RP	12.85030	4	3.21258	1.67	0.1595
2 ERROR	298.05556	155	1.92294		
SCAT	164.00030	1	164.00030	28.04	0.0000
SP	9.08207	4	2.27052	0.39	0.8169
3 ERROR	906.69878	155	5.84967		
RS	34.59609	1	34.59609	19.22	0.0000
RSP	12.47270	4	3.11817	1.73	0.1455
4 ERROR	278.96441	155	1.79977		
WO	1349.98522	3	449.99507	179.12	0.0000
WP	33.74466	12	2.81206	1.12	0.3418
5 ERROR	1168.19792	465	2.51225		
RW	14.40280	3	4.80093	4.73	0.0029
RWP	19.22591	12	1.60216	1.58	0.0946
6 ERROR	472.22917	465	1.01555		
SW	125.83163	3	41.94388	30.81	0.0000
SWP	115.61567	12	9.63464	7.08	0.0000
7 ERROR	633.02691	465	1.36135		
RSW	17.66927	3	5.88976	5.75	0.0007
RSWP	37.79379	12	3.14948	3.08	0.0003
8 ERROR	475.98003	465	1.02361		

TABLE B.40: Card-Sorting. Restructuring verbs, all levels, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	377.11054	1	377.11054	80.77	0.0000
PLEVEL	64.77951	4	16.19488	3.47	0.0096
1 ERROR	723.65799	155	4.66876		
RC	33.84696	1	33.84696	15.36	0.0001
RP	33.12639	4	8.28160	3.76	0.0060
2 ERROR	341.47049	155	2.20304		
WO	215.20822	3	71.73607	13.69	0.0000
WP	128.80035	12	10.73336	2.05	0.0190
3 ERROR	2435.73090	465	5.23813		
RW	68.32235	3	22.77412	10.13	0.0000
RWP	33.91597	12	2.82633	1.26	0.2409
4 ERROR	1045.10590	465	2.24754		

TABLE B.41: Card-Sorting. Unergative verbs, native vs near-native, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	9645.93508	1	9645.93508	552.65	0.0000
NLG	96.60833	2	48.30417	2.77	0.0691
1 ERROR	1343.95833	77	17.45400		
SCAT	79.31321	2	39.65661	30.38	0.0000
SN	6.04306	4	1.51076	1.16	0.3319
2 ERROR	201.01528	154	1.30529		
WO	283.58400	1	283.58400	70.70	0.0000
WN	65.04537	2	32.52269	8.11	0.0006
3 ERROR	308.85463	77	4.01110		
SW	0.00744	2	0.00372	0.00	0.9979
SWN	8.11991	4	2.02998	1.16	0.3328
4 ERROR	270.58009	154	1.75701		
AUX	690.55919	1	690.55919	128.03	0.0000
AN	32.02870	2	16.01435	2.97	0.0573
5 ERROR	415.32130	77	5.39378		
SA	404.32269	2	202.16134	61.77	0.0000
SAN	14.37546	4	3.59387	1.10	0.3596
6 ERROR	504.02454	154	3.27289		
WA	321.33439	1	321.33439	149.66	0.0000
WAN	41.36204	2	20.68102	9.63	0.0002
7 ERROR	165.32130	77	2.14703		
SWA	9.68651	2	4.84325	4.15	0.0176
SWAN	1.16157	4	0.29039	0.25	0.9102
8 ERROR	179.93009	154	1.16838		

TABLE B.42: Card-Sorting. Unergative verbs, native vs near-native, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	1381.11839	1	1381.11839	128.03	0.0000
NLG	64.05741	2	32.02870	2.97	0.0573
1 ERROR	830.64259	77	0.78757		
SCAT	808.64537	2	404.32269	61.77	0.0000
SN	28.75093	4	7.18773	1.10	0.3596
2 ERROR	1008.04907	154	6.54577		
WO	642.66878	1	642.66878	149.66	0.0000
WN	82.72407	2	41.36204	9.63	0.0002
3 ERROR	330.64259	77	4.29406		
SW	19.37302	2	9.68651	4.15	0.0176
SWN	2.32315	4	0.58079	0.25	0.9102
4 ERROR	359.86019	154	2.33675		

TABLE B.43: Card-Sorting. Unaccusative verbs, native vs near-native, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	19832.47055	1	19832.47055	681.59	0.0000
NLG	75.40347	2	37.70174	1.30	0.2796
1 ERROR	2240.48653	77	29.09723		
SCAT	99.49856	4	24.87464	16.59	0.0000
SN	7.01181	8	0.87648	0.58	0.7905
2 ERROR	461.79194	308	1.49932		
WO	239.61166	1	239.61166	60.75	0.0000
WN	26.20208	2	13.10104	3.32	0.0413
3 ERROR	303.69542	77	3.94410		
SW	7.89098	4	1.97274	1.78	0.1330
SWN	11.28819	8	1.41102	1.27	0.2575
4 ERROR	341.65806	308	1.10928		
AUX	2428.29962	1	2428.29962	189.48	0.0000
AN	174.51347	2	87.25674	6.81	0.0019
5 ERROR	986.78653	77	12.81541		
SA	250.58461	4	62.64615	25.69	0.0000
SAN	15.77681	8	1.97210	0.81	0.5953
6 ERROR	751.04194	308	2.43845		
WA	19.30887	1	19.30887	10.16	0.0021
WAN	12.39208	2	6.19604	3.26	0.0438
7 ERROR	146.38542	77	1.90111		
SWA	30.43059	4	7.60765	6.28	0.0001
SWAN	13.31486	8	1.66436	1.37	0.2076
8 ERROR	373.35139	308	1.21218		

TABLE B.44: Card-Sorting. Unaccusative verbs, native vs near-native, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	4856.59923	1	4856.59923	189.48	0.0000
NLG	349.02694	2	174.51347	6.81	0.0019
1 ERROR	1973.57306	77	25.63082		
SCAT	501.16921	4	125.29230	25.69	0.0000
SN	31.55361	8	3.94420	0.81	0.5953
2 ERROR	1502.08389	308	4.87690		
AUX	38.61773	1	38.61773	10.16	0.0021
AN	24.78417	2	12.39208	3.26	0.0438
3 ERROR	292.77083	77	3.80222		
SA	60.86118	4	15.21529	6.28	0.0001
SAN	26.62972	8	3.32872	1.37	0.2076
4 ERROR	746.70278	308	2.42436		

TABLE B.45: Card-Sorting. Restructuring verbs, native vs near-native, all variables (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	16741.19206	1	16741.19206	510.37	0.0000
NLG	91.35851	2	45.67925	1.39	0.2546
1 ERROR	2525.77899	77	32.80232		
RC	2.19303	1	2.19303	0.96	0.3312
RN	0.70851	2	0.35425	0.15	0.8571
2 ERROR	176.59149	77	2.29340		
SCAT	30.46677	1	30.46677	4.83	0.0310
SN	9.72622	2	4.86311	0.77	0.4662
3 ERROR	485.82066	77	6.30936		
RS	18.16312	1	18.16312	8.95	0.0037
RSN	5.29844	2	2.64922	1.31	0.2768
4 ERROR	156.19844	77	2.02855		
WO	787.56833	3	262.52278	98.59	0.0000
WN	14.39635	6	2.39939	0.90	0.4948
5 ERROR	615.11615	231	2.66284		
RW	21.93932	3	7.31311	6.56	0.0003
RWN	4.58247	6	0.76374	0.68	0.6620
6 ERROR	257.59253	231	1.11512		
SW	176.78457	3	58.92819	40.36	0.0000
SWN	25.14392	6	4.19065	2.87	0.0102
7 ERROR	337.29670	231	1.46016		
RSW	37.74326	3	12.58109	10.12	0.0000
RSWN	7.27170	6	1.21195	0.97	0.4429
8 ERROR	287.16892	231	1.24316		

TABLE B.46: Card-Sorting. Restructuring verbs, native vs near-native, auxiliary preferences (BY SUBJECTS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	235.63957	1	235.63957	45.24	0.0000
NLG	41.89358	2	20.94679	4.02	0.0218
1 ERROR	401.10486	77	5.20915		
RC	60.18169	1	60.18169	25.04	0.0000
RN	1.96719	2	0.98359	0.41	0.6656
2 ERROR	185.09375	77	2.40381		
WO	47.35468	3	15.78489	2.64	0.0502
WN	73.77795	6	12.29633	2.06	0.0593
3 ERROR	1381.16736	231	5.97908		
RW	39.74474	3	13.24825	5.65	0.0009
RWN	17.55434	6	2.92572	1.25	0.2831
4 ERROR	541.92847	231	2.34601		

B.2.2.2 BY MATERIALS

TABLE B.47: Card-Sorting. Unergative verbs, all levels, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	2764.02089	1	2764.02089	4047.02	0.0000
SCAT	1.18939	2	0.59469	0.87	0.4355
WO	34.12602	1	34.12602	49.97	0.0000
SW	0.29673	2	0.14837	0.22	0.8068
1 ERROR	12.29358	18	0.68298		
PLEVEL	13.46040	4	3.36510	4.72	0.0019
PS	8.50350	8	1.06294	1.49	0.1759
PW	13.35939	4	3.33985	4.68	0.0020
PSW	0.40418	8	0.05052	0.07	0.9998
2 ERROR	51.35542	72	0.71327		
AUX	93.78472	1	93.78472	142.79	0.0000
AS	51.23440	2	25.61720	39.00	0.0000
AW	38.71176	1	38.71176	58.94	0.0000
ASW	1.02236	2	0.51118	0.78	0.4740
3 ERROR	11.82235	18	0.65680		
PA	26.08339	4	6.52085	14.69	0.0000
PAS	16.84797	8	2.10600	4.74	0.0001
PAW	13.20792	4	3.30198	7.44	0.0000
PASW	1.27048	8	0.15881	0.36	0.9392
4 ERROR	31.96481	72	0.44396		

TABLE B.48: Card-Sorting. Unergative verbs, all levels, auxiliary preferences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	187.56943	1	187.56943	142.79	0.0000
SCAT	102.46879	2	51.23440	39.00	0.0000
WO	77.42352	1	77.42352	58.94	0.0000
SW	2.04472	2	1.02236	0.78	0.4740
1 ERROR	23.64470	18	1.31359		
PLEVEL	52.16679	4	13.04170	14.69	0.0000
PS	33.69594	8	4.21199	4.74	0.0001
PW	26.41584	4	6.60396	7.44	0.0000
PSW	2.54097	8	0.31762	0.36	0.9392
2 ERROR	63.92962	72	0.88791		

TABLE B.49: Card-Sorting. Unaccusative verbs, all levels, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	4974.16740	1	4974.16740	4911.15	0.0000
SCAT	7.93883	4	1.98471	1.96	0.1263
WO	56.98152	1	56.98152	56.26	0.0000
SW	1.78197	4	0.44549	0.44	0.7788
1 ERROR	30.38494	30	1.01283		
PLEVEL	26.28757	4	6.57189	9.10	0.0000
PS	11.67295	16	0.72956	1.01	0.4505
PW	2.96639	4	0.74160	1.03	0.3961
PSW	2.50680	16	0.15668	0.22	0.9994
2 ERROR	86.62370	120	0.72186		
AUX	290.51323	1	290.51323	151.52	0.0000
AS	53.18482	4	13.29620	6.93	0.0004
AW	0.17710	1	0.17710	0.09	0.7633
ASW	4.38450	4	1.09612	0.57	0.6852
3 ERROR	57.52109	30	1.91737		
PA	102.04052	4	25.51013	53.94	0.0000
PAS	5.67522	16	0.35470	0.75	0.7376
PAW	4.01606	4	1.00401	2.12	0.0821
PASW	3.25960	16	0.20372	0.43	0.9716
4 ERROR	56.74822	120	0.47290		

TABLE B.50: Card-Sorting. Unaccusative verbs, all levels, auxiliary preferences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	581.02645	1	581.02645	151.52	0.0000
SCAT	106.36964	4	26.59241	6.93	0.0004
WO	0.35421	1	0.35421	0.09	0.7633
SW	8.76899	4	2.19225	0.57	0.6852
1 ERROR	115.04218	30	3.83474		
PLEVEL	204.08103	4	51.02026	53.94	0.0000
PS	11.35043	16	0.70940	0.75	0.7376
PW	8.03211	4	2.00803	2.12	0.0821
PSW	6.51920	16	0.40745	0.43	0.9716
2 ERROR	113.49644	120	0.94580		

TABLE B.51: Card-Sorting. Restructuring verbs, all levels, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	3529.86123	1	3529.86123	5752.82	0.0000
RC	0.05944	1	0.05944	0.10	0.7583
WO	172.49817	3	57.49939	93.71	0.0000
RW	1.84036	3	0.61345	1.00	0.4099
1 ERROR	14.72611	24	0.61359		
PLEVEL	65.64129	4	16.41032	21.07	0.7014
PW	3.99091	12	0.33258	0.43	0.9492
PRW	2.62568	12	0.21881	0.28	0.9911
2 ERROR	74.77940	96	0.77895		
SCAT	20.95558	1	20.95558	41.19	0.0000
SR	4.42060	1	4.42060	8.69	0.0070
SW	16.07851	3	5.35950	10.54	0.0001
SRW	2.25774	3	0.75258	1.48	0.2453
3 ERROR	12.20934	24	0.50872		
PS	1.04199	4	0.26050	0.83	0.5072
PSR	1.70552	4	0.42638	1.36	0.2522
PSW	13.44082	12	1.12007	3.58	0.0002
PSRW	4.82672	12	0.40223	1.29	0.2389
4 ERROR	30.00860	96	0.31259		

TABLE B.52: Card-Sorting. Restructuring verbs, all levels, auxiliary preferences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	55.19242	1	55.19242	54.25	0.0000
RC	7.57139	1	7.57139	7.44	0.0117
WO	18.87576	3	6.29192	6.18	0.0029
RW	5.78529	3	1.92843	1.90	0.1573
1 ERROR	24.41867	24	1.01744		
PLEVEL	10.89593	4	2.72398	4.36	0.0028
PR	7.84700	4	1.96175	3.14	0.0180
PW	18.06971	12	1.50581	2.41	0.0090
PRW	5.21748	12	0.43479	0.70	0.7522
2 ERROR	60.01719	96	0.62518		

TABLE B.53: Card-Sorting. Unergative verbs, native vs near-native, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	1536.20451	1	1536.20451	867.31	0.0000
SCAT	12.63136	2	6.31568	3.57	0.0496
WO	45.16335	1	45.16335	25.50	0.0001
SW	0.00118	2	0.00059	0.00	0.9997
1 ERROR	31.88228	18	1.77124		
NLG	16.77007	2	8.38504	11.08	0.0002
NS	0.98402	4	0.24601	0.33	0.8593
NW	9.19898	2	4.59949	6.08	0.0053
NSW	1.43292	4	0.35823	0.47	0.7549
2 ERROR	27.23864	36	0.75663		
AUX	109.97797	1	109.97797	129.08	0.0000
AS	64.39213	2	32.19606	37.79	0.0000
AW	51.17549	1	51.17549	60.06	0.0000
ASW	1.54267	2	0.77133	0.91	0.4221
3 ERROR	15.33628	18	0.85202		
NA	5.84417	2	2.92208	3.95	0.0282
NAS	2.25433	4	0.56358	0.76	0.5575
NAW	5.55774	2	2.77887	3.75	0.0331
NASW	0.15725	4	0.03931	0.05	0.9945
4 ERROR	26.65777	36	0.74049		

TABLE B.54: Card-Sorting. Unergative verbs, native vs near-native, auxiliary preferences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	219.95595	1	219.95595	129.08	0.0000
SCAT	128.78426	2	64.39213	37.79	0.0000
WO	102.35099	1	102.35099	60.06	0.0000
SW	3.08534	2	1.54267	0.91	0.4221
1 ERROR	30.67256	18	1.70403		
NLG	11.68833	2	5.84417	3.95	0.0282
NS	4.50865	4	1.12716	0.76	0.5575
NW	11.11547	2	5.55774	3.75	0.0331
NSW	0.31449	4	0.07862	0.05	0.9945
2 ERROR	53.31553	36	1.48099		

TABLE B.55: Card-Sorting. Unaccusative verbs, native vs near-native, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	3158.50445	1	3158.50445	2293.96	0.0000
SCAT	15.84606	4	3.96152	2.88	0.0396
WO	38.16036	1	38.16036	27.72	0.0000
SW	1.25671	4	0.31418	0.23	0.9204
1 ERROR	41.30631	30	1.37688		
NLG	13.96183	2	6.98091	9.31	0.0003
NS	1.02094	8	0.12762	0.17	0.9941
NW	4.44769	2	2.22384	2.97	0.0591
NSW	1.67907	8	0.20988	0.28	0.9701
2 ERROR	44.98675	60	0.74978		
AUX	386.72931	1	386.72931	312.78	0.0000
AS	39.90798	4	9.97700	8.07	0.0002
AW	3.07513	1	3.07513	2.49	0.1253
ASW	4.84635	4	1.21159	0.98	0.4333
3 ERROR	37.09286	30	1.23643		
NA	24.11106	2	12.05553	17.56	0.0000
NAS	2.37148	8	0.29643	0.43	0.8974
NAW	1.70163	2	0.85081	1.24	0.2969
NASW	1.85891	8	0.23236	0.34	0.9475
4 ERROR	41.19310	60	0.68655		

TABLE B.56: Card-Sorting. Unaccusative verbs, native vs near-native, auxiliary preferences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	773.45862	1	773.45862	312.78	0.0000
SCAT	79.81597	4	19.95399	8.07	0.0002
WO	6.15026	1	6.15026	2.49	0.1253
SW	9.69270	4	2.42318	0.98	0.4333
1 ERROR	74.18572	30	2.47286		
NLG	48.22212	2	24.11106	17.56	0.0000
NS	4.74296	8	0.59287	0.43	0.8974
NW	3.40326	2	1.70163	1.24	0.2969
NSW	3.71783	8	0.46473	0.34	0.9475
2 ERROR	82.38619	60	1.37310		

TABLE B.57: Card-Sorting. Restructuring verbs, native vs near-native, all variables (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	2658.74217	1	2658.74217	2949.52	0.0000
RC	0.33241	1	0.33241	0.37	0.5494
WO	125.52987	3	41.84329	46.42	0.0000
RW	3.40607	3	1.13536	1.26	0.3105
1 ERROR	21.63396	24	0.90142		
NLG	12.87553	2	6.43777	5.19	0.0091
NR	0.14028	2	0.07014	0.06	0.9451
NW	2.36111	6	0.39352	0.32	0.9248
NRW	0.84780	6	0.14130	0.11	0.9943
2 ERROR	59.53084	48	1.24023		
SCAT	5.17526	1	5.17526	6.69	0.0162
SR	2.84375	1	2.84375	3.68	0.0671
SW	26.88423	3	8.96141	11.59	0.0001
SRW	6.12096	3	2.04032	2.64	0.0726
3 ERROR	18.55560	24	0.77315		
NS	1.54752	2	0.77376	3.27	0.0467
NSR	0.99793	2	0.49896	2.11	0.1327
NSW	3.55663	6	0.59277	2.50	0.0345
NSRW	1.01855	6	0.16976	0.72	0.6378
4 ERROR	11.36486	48	0.23677		

TABLE B.58: Card-Sorting. Restructuring verbs, native vs near-native, auxiliary preferences (BY MATERIALS)

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F	TAIL PROB.
MEAN	60.14912	1	60.14912	38.90	0.0000
RC	15.22693	1	15.22693	9.85	0.0045
WO	3.96986	3	1.32329	0.86	0.4774
RW	2.70250	3	0.90083	0.58	0.6322
1 ERROR	37.11120	24	1.54630		
NLG	2.76066	2	1.38033	2.91	0.0639
NR	1.34974	2	0.67487	1.43	0.2505
NW	7.44762	6	1.24127	2.62	0.0280
NRW	2.68321	6	0.44720	0.94	0.4725
2 ERROR	22.72973	48	0.47354		

APPENDIX C

CARD-SORTING: RESULTS

TABLE C.1: Mean acceptability judgments on **unergative** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

	BEG	INT	ADV	ENNS	INS
NON-MOTIONAL, BASIC, ESSERE	3.125	2.361	2.062	1.625	1.167
NON-MOTIONAL, BASIC, AVERE	4.844	4.500	4.812	5.750	6.139
NON-MOTIONAL, NE-CL, ESSERE	3.062	2.778	2.281	2.167	1.250
NON-MOTIONAL, NE-CL, AVERE	3.625	3.278	3.719	4.667	2.889
MOTIONAL, BASIC, ESSERE	3.875	2.972	2.750	2.042	1.278
MOTIONAL, BASIC, AVERE	4.625	3.972	4.281	5.750	6.222
MOTIONAL, NE-CL, ESSERE	3.312	3.028	2.938	2.500	1.778
MOTIONAL, NE-CL, AVERE	3.688	2.917	3.156	4.333	2.556
[+UNACCUSATIVE ALTERNANT], BASIC, ESSERE	3.594	3.083	3.281	3.917	4.028
[+UNACCUSATIVE ALTERNANT], BASIC, AVERE	3.688	3.917	3.656	4.667	5.139
[+UNACCUSATIVE ALTERNANT], NE-CL, ESSERE	3.438	2.722	3.281	3.833	3.528
[+UNACCUSATIVE ALTERNANT], NE-CL, AVERE	2.656	2.944	2.812	3.500	1.861

TABLE C.2: Mean auxiliary preferences for **unergative** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

	BEG	INT	ADV	ENNS	INS
NON-MOTIONAL, BASIC	1.719	2.139	2.750	4.125	4.972
NON-MOTIONAL, NE-CLITICIZATION	0.562	0.500	1.438	2.500	1.639
MOTIONAL, BASIC	0.750	1.000	1.531	3.708	4.944
MOTIONAL, NE-CLITICIZATION	0.375	-0.111	0.219	1.833	0.778
[+UNACCUSATIVE ALTERNANT], BASIC	0.094	0.833	0.375	0.750	1.111
[+UNACCUSATIVE ALTERNANT], NE-CLITIC.	-0.781	0.222	-0.469	-0.333	-1.667

FIGURE C.1: Mean acceptability judgments on **unergative** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

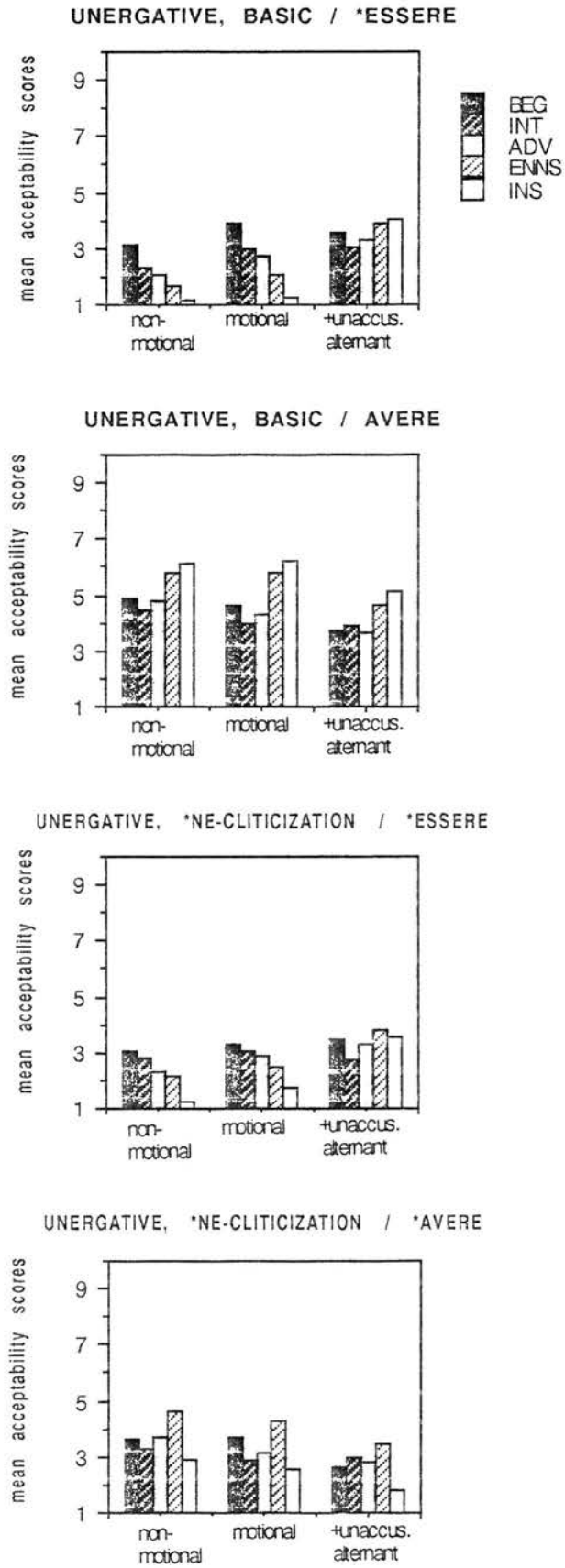
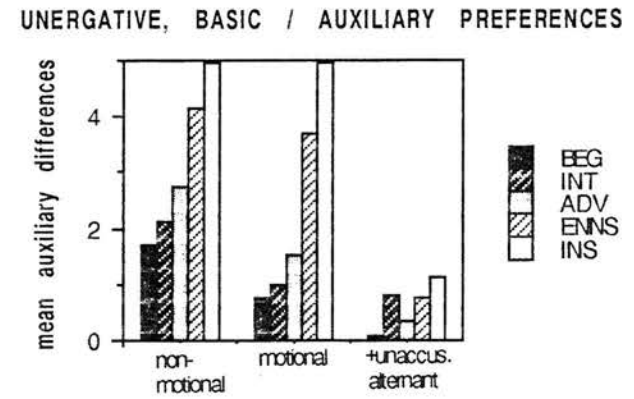


FIGURE C.2: Mean auxiliary preferences for **unergative** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)



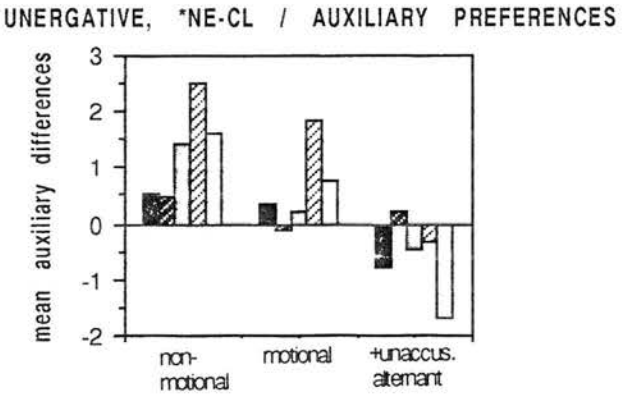


TABLE C.3: Mean acceptability judgments on **unaccusative** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

	BEG	INT	ADV	ENNS	INS
CHANGE-OF-LOCATION, BASIC, ESSERE	5.094	4.333	4.688	5.750	6.250
CHANGE-OF-LOCATION, BASIC, AVERE	3.219	2.722	2.219	1.708	1.167
CHANGE-OF-LOCATION, NE-CL, ESSERE	4.156	3.556	3.969	5.292	5.028
CHANGE-OF-LOCATION, NE-CL, AVERE	1.969	1.694	1.406	1.167	0.944
CONTINUATION-OF-STATE, BASIC, ESSERE	4.500	3.694	4.500	5.042	5.889
CONTINUATION-OF-STATE, BASIC, AVERE	3.531	3.361	2.938	3.042	2.111
CONTINUATION-OF-STATE, NE-CL, ESSERE	4.062	3.389	4.125	4.750	4.611
CONTINUATION-OF-STATE, NE-CL, AVERE	2.500	2.111	1.625	2.250	1.139
EXISTENCE-OF-STATE, BASIC, ESSERE	4.312	3.917	4.219	4.792	6.083
EXISTENCE-OF-STATE, BASIC, AVERE	3.875	3.667	3.281	3.250	2.278
EXISTENCE-OF-STATE, NE-CL, ESSERE	3.375	3.167	3.344	4.667	4.500
EXISTENCE-OF-STATE, NE-CL, AVERE	2.719	2.778	2.250	2.750	1.306
[+TRANSITIVE ALTERNANT], BASIC, ESSERE	4.125	4.056	4.062	5.208	6.222
[+TRANSITIVE ALTERNANT], BASIC, AVERE	4.125	3.333	3.469	3.083	2.222
[+TRANSITIVE ALTERNANT], NE-CL, ESSERE	3.750	2.917	3.312	4.667	4.639
[+TRANSITIVE ALTERNANT], NE-CL, AVERE	3.375	2.694	3.344	3.167	3.083
[+UNERGATIVE ALTERNANT], BASIC, ESSERE	3.750	3.722	3.969	5.333	5.917
[+UNERGATIVE ALTERNANT], BASIC, AVERE	3.344	3.472	3.250	3.792	3.306
[+UNERGATIVE ALTERNANT], NE-CL, ESSERE	2.938	2.917	3.000	4.542	4.833
[+UNERGATIVE ALTERNANT], NE-CL, AVERE	2.906	2.750	2.281	3.542	2.194

TABLE C.4: Mean auxiliary preferences for **unaccusative** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

	BEG	INT	ADV	ENNS	INS
CHANGE-OF-LOCATION, BASIC	1.875	1.611	2.469	4.042	5.083
CHANGE-OF-LOCATION, NE-CLITICIZATION	2.188	1.861	2.562	4.125	4.083
CONTINUATION-OF-STATE, BASIC	0.969	0.333	1.562	2.000	3.778
CONTINUATION-OF-STATE, NE-CLITICIZATION	1.562	1.278	2.500	2.500	3.472
EXISTENCE-OF-STATE, BASIC	0.438	0.250	0.938	1.542	3.806
EXISTENCE-OF-STATE, NE-CLITICIZATION	0.656	0.389	1.094	1.917	3.194
[+TRANSITIVE ALTERNANT], BASIC	0.000	0.722	0.594	2.125	4.000
[+TRANSITIVE ALTERNANT], NE-CLITICIZATION	0.375	0.222	-0.031	1.500	1.556
[+UNACCUSATIVE ALTERNANT], BASIC	0.406	0.250	0.719	1.542	2.611
[+UNACCUSATIVE ALTERNANT], NE-CLIT.	0.031	0.167	0.719	1.000	2.639

FIGURE C.3: Mean acceptability judgments on **unaccusative** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

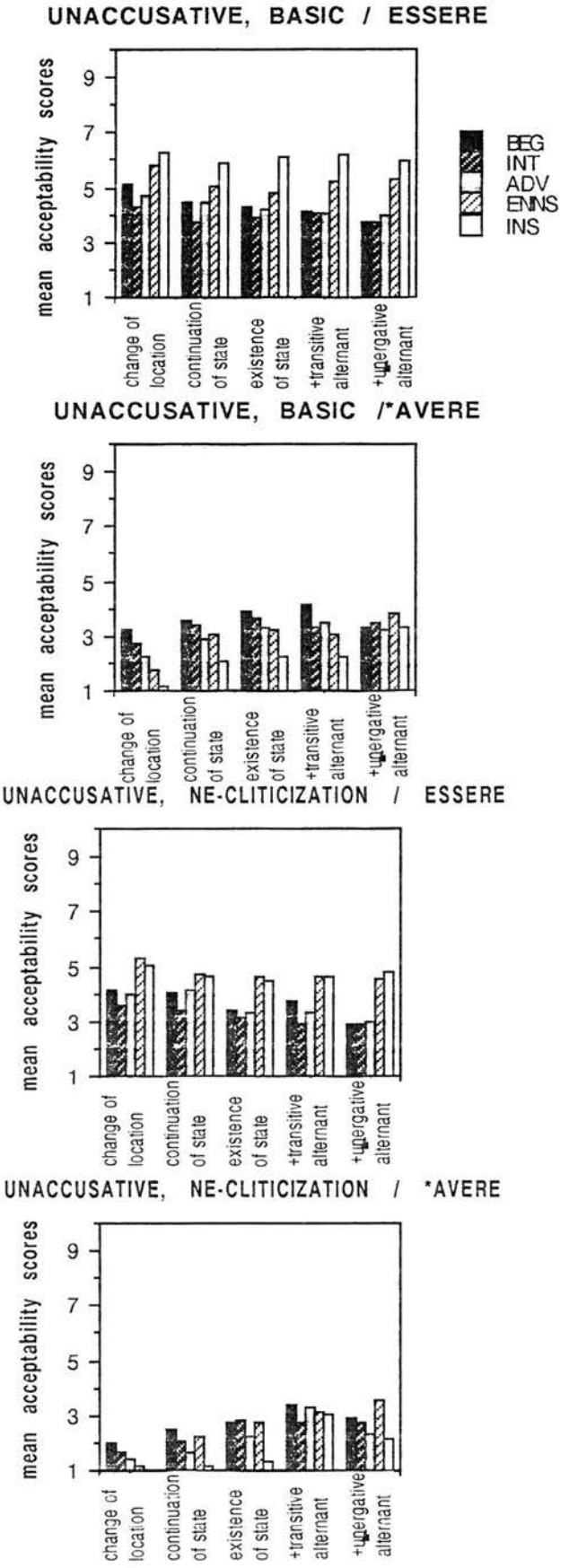
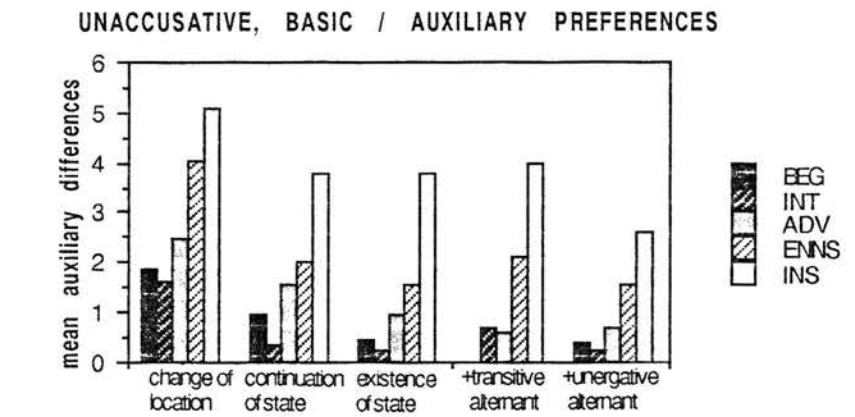


FIGURE C.4: Mean auxiliary preferences for **unaccusative** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)



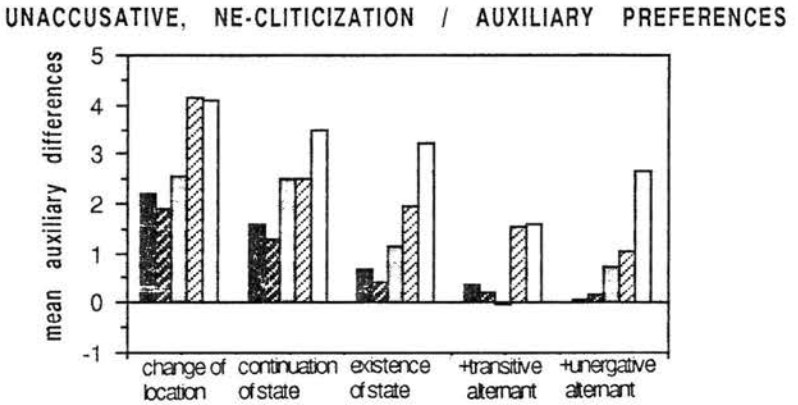


TABLE C.5: Means of acceptability judgments on **restructuring** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

	BEG	INT	ADV	ENNS	INS
RAISING, BASIC, NON-RESTRUCTURED	4.656	4.167	4.500	5.042	5.167
RAISING, CLEFTING, NON-RESTRUCTURED	2.188	1.833	1.875	2.417	3.333
RAISING, CLIT-CLIMB, NON-RESTRUCTURED	3.031	3.111	2.781	3.792	3.750
RAISING, NO-CLIT-MOV, NON-RESTRUCTURED	3.188	3.139	2.938	3.625	4.389
RAISING, BASIC, RESTRUCTURED	3.781	3.139	3.531	4.042	4.944
RAISING, CLEFTING, RESTRUCTURED	1.625	2.000	1.625	2.083	2.444
RAISING, CLIT-CLIMB, RESTRUCTURED	2.844	2.917	2.844	4.125	4.667
RAISING, NO-CLIT-MOV, RESTRUCTURED	2.844	2.667	2.875	4.083	4.306
CONTROL, BASIC, NON-RESTRUCTURING	4.781	4.222	4.656	5.375	5.778
CONTROL, CLEFTING, NON-RESTRUCTURING	2.000	2.222	2.094	2.958	3.833
CONTROL, CLIT-CLIMB, NON-RESTRUCTURING	3.250	3.250	3.531	2.875	2.778
CONTROL, NO-CLIT-MOV, NON-RESTRUCTURING	3.656	3.361	3.906	4.750	4.889
CONTROL, BASIC, RESTRUCTURING	3.688	2.778	3.531	3.375	4.417
CONTROL, CLEFTING, RESTRUCTURING	2.062	1.917	1.562	2.167	2.194
CONTROL, CLIT-CLIMB, RESTRUCTURING	2.969	2.528	2.656	3.500	4.611
CONTROL, NO-CLIT-MOV, RESTRUCTURING	3.188	2.444	2.781	3.000	3.861

TABLE C.6: Mean auxiliary preferences for **restructuring** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

	BEG	INT	ADV	ENNS	INS
RAISING, BASIC	0.875	1.028	0.969	1.000	0.222
RAISING, CLEFTING	0.562	-0.167	0.250	0.333	0.889
RAISING, CLITIC-CLIMBING	-0.188	-0.194	0.062	0.125	0.917
RAISING, NO-CLITIC-MOVEMENT	0.344	0.472	0.062	0.042	0.083
CONTROL, BASIC	1.094	1.444	1.125	2.000	1.361
CONTROL, CLEFTING	-0.062	0.306	0.531	0.792	1.639
CONTROL, CLITIC-CLIMBING	-0.281	-0.722	-0.875	-0.542	1.167
CONTROL, NO-CLITIC-MOVEMENT	0.469	0.917	1.125	1.750	1.028

FIGURE C.5: Means of acceptability judgments on **Raising** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)

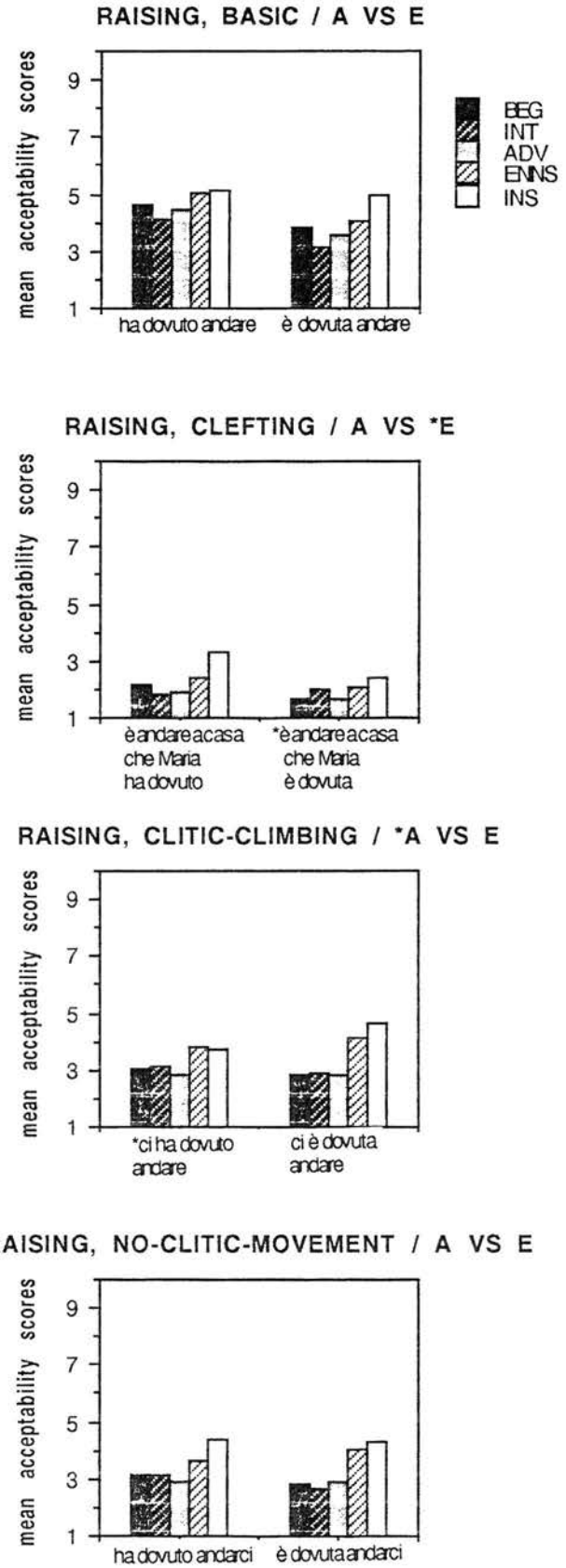
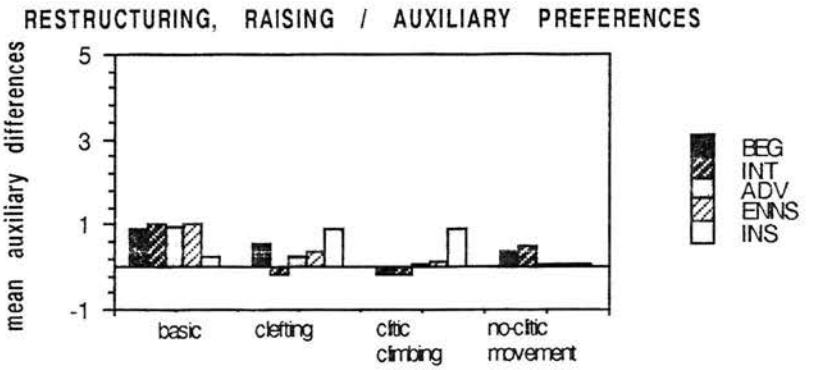


FIGURE C.7: Mean auxiliary preferences for **restructuring (Raising and Control)** verbs (BEG = beginners, INT = intermediate, ADV = advanced, ENNS = near-native speakers, INS = native speakers)



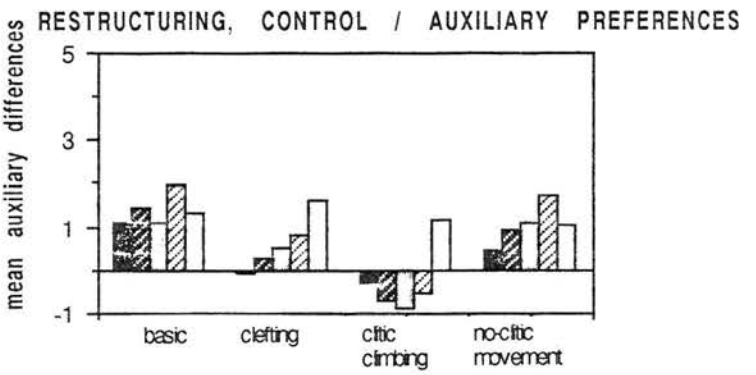


TABLE C.7: Mean acceptability judgments on **unergative** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).

	ENNS	FNNS	INS
NON-MOTIONAL, BASIC, ESSERE	1.625	1.600	1.167
NON-MOTIONAL, BASIC, AVERE	5.750	5.150	6.139
NON-MOTIONAL, NE-CL, ESSERE	2.167	1.250	1.250
NON-MOTIONAL, NE-CL, AVERE	4.667	2.450	2.889
MOTIONAL, BASIC, ESSERE	2.042	2.000	1.278
MOTIONAL, BASIC, AVERE	5.750	4.700	6.222
MOTIONAL, NE-CL, ESSERE	2.500	2.000	1.778
MOTIONAL, NE-CL, AVERE	4.333	2.150	2.556
[+UNACCUSATIVE ALTERNANT], BASIC, ESSERE	3.917	3.550	4.028
[+UNACCUSATIVE ALTERNANT], NE-CL, AVERE	4.667	4.150	5.139
[+UNACCUSATIVE ALTERNANT], BASIC, ESSERE	3.833	3.550	3.528
[+UNACCUSATIVE ALTERNANT], NE-CL, AVERE	3.500	2.450	1.861

TABLE C.8: Mean auxiliary preferences for **unergative** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).

	ENNS	FNNS	INS
NON-MOTIONAL, BASIC	4.125	3.550	4.972
NON-MOTIONAL, NE-CLITICIZATION	2.500	1.200	1.639
MOTIONAL, BASIC	3.708	2.700	4.944
MOTIONAL, NE-CLITICIZATION	1.833	0.150	0.778
[+UNACCUSATIVE ALTERNANT], BASIC	0.750	0.600	1.111
[+UNACCUSATIVE ALTERNANT], NE-CL	-0.333	-1.100	-1.667

FIGURE C.8: Mean acceptability judgments on **unergative** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).

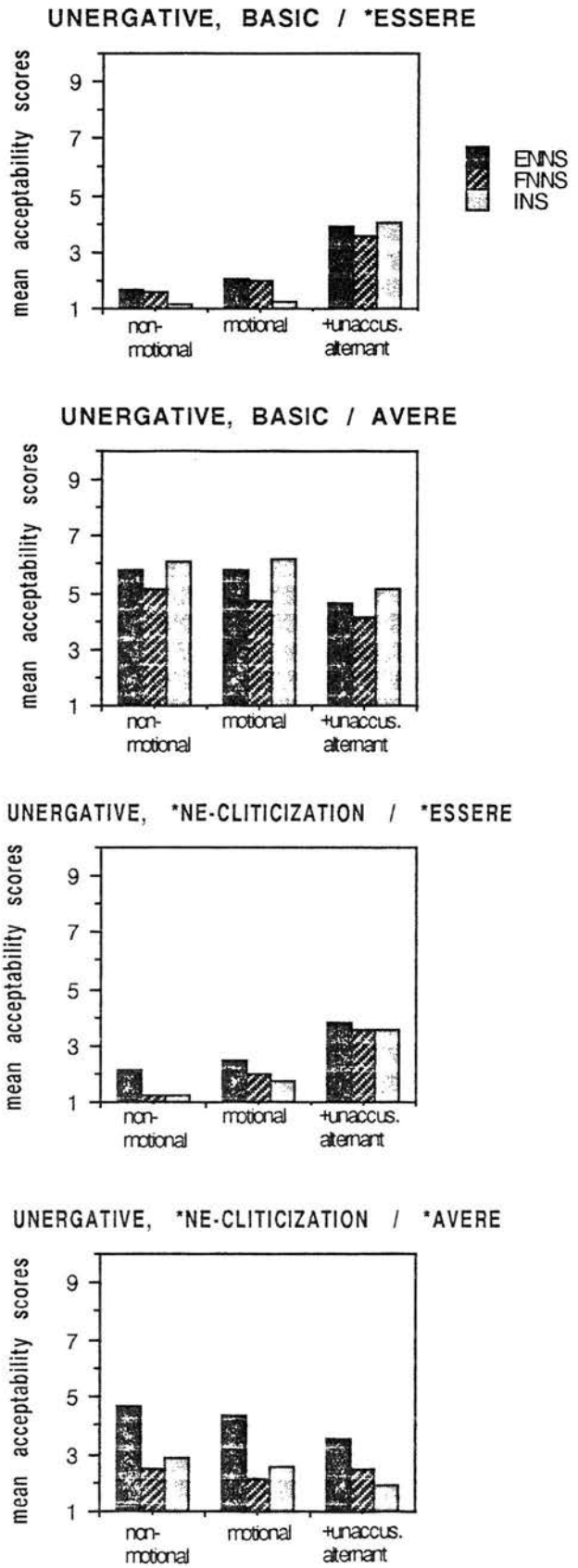
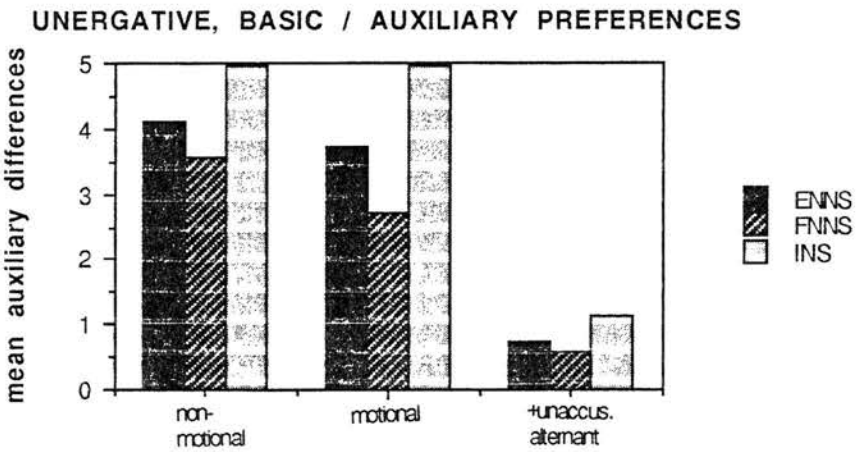


FIGURE C.9: Mean auxiliary preferences for **unergative** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).



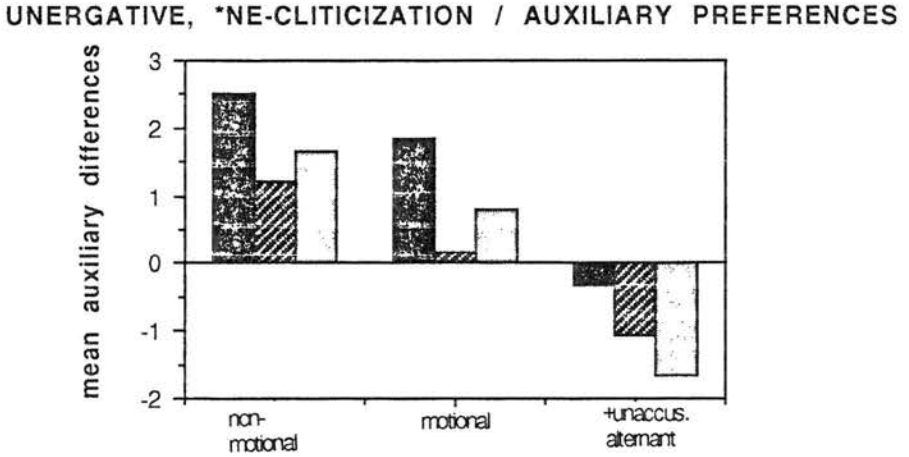


TABLE C.9: Mean acceptability judgments on **unaccusative** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).

	ENNS	FNNS	INS
CHANGE-OF-LOCATION, BASIC, ESSERE	5.750	5.050	6.250
CHANGE-OF-LOCATION, BASIC, AVERE	1.708	1.800	1.167
CHANGE-OF-LOCATION, NE-CL, ESSERE	5.292	4.250	5.028
CHANGE-OF-LOCATION, NE-CL, AVERE	1.167	0.900	0.944
CONTINUATION-OF-STATE, BASIC, ESSERE	5.042	4.950	5.889
CONTINUATION-OF-STATE, BASIC, AVERE	3.042	2.600	2.111
CONTINUATION-OF-STATE, NE-CL, ESSERE	4.750	3.900	4.611
CONTINUATION-OF-STATE, NE-CL, AVERE	2.250	1.350	1.139
EXISTENCE-OF-STATE, BASIC, ESSERE	4.792	4.850	6.083
EXISTENCE-OF-STATE, BASIC, AVERE	3.250	2.900	2.278
EXISTENCE-OF-STATE, NE-CL, ESSERE	4.667	3.400	4.500
EXISTENCE-OF-STATE, NE-CL, AVERE	2.750	1.900	1.306
[+TRANSITIVE ALTERNANT], BASIC, ESSERE	5.208	5.100	6.222
[+TRANSITIVE ALTERNANT], BASIC, AVERE	3.083	3.050	2.222
[+TRANSITIVE ALTERNANT], NE-CL, ESSERE	4.667	3.200	4.639
[+TRANSITIVE ALTERNANT], NE-CL, AVERE	3.167	2.700	3.083
[+UNERGATIVE ALTERNANT], BASIC, ESSERE	5.333	4.700	5.917
[+UNERGATIVE ALTERNANT], BASIC, AVERE	3.792	3.350	3.306
[+UNERGATIVE ALTERNANT], NE-CL, ESSERE	4.542	3.500	4.833
[+UNERGATIVE ALTERNANT], NE-CL, AVERE	3.542	2.700	2.194

TABLE C.10: Mean auxiliary preferences for **unaccusative** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).

	ENNS	FNNS	INS
CHANGE-OF-LOCATION, BASIC	4.042	3.250	5.083
CHANGE-OF-LOCATION, NE-CLITICIZATION	4.125	3.350	4.083
CONTINUATION-OF-STATE, BASIC	2.000	2.350	3.778
CONTINUATION-OF-STATE, NE-CLITICIZATION	2.500	2.550	3.472
EXISTENCE-OF-STATE, BASIC	1.542	1.950	3.806
EXISTENCE-OF-STATE, NE-CLITICIZATION	1.917	1.500	3.194
[+TRANSITIVE ALTERNANT], BASIC	2.125	2.050	4.000
[+TRANSITIVE ALTERNANT], NE-CLITICIZATION	1.500	0.500	1.556
[+UNACCUSATIVE ALTERNANT], BASIC	1.542	1.350	2.611
[+UNACCUSATIVE ALTERNANT], NE-CLITICIZATION	1.000	0.800	2.639

FIGURE C.10: Mean acceptability judgments on **unaccusative** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).

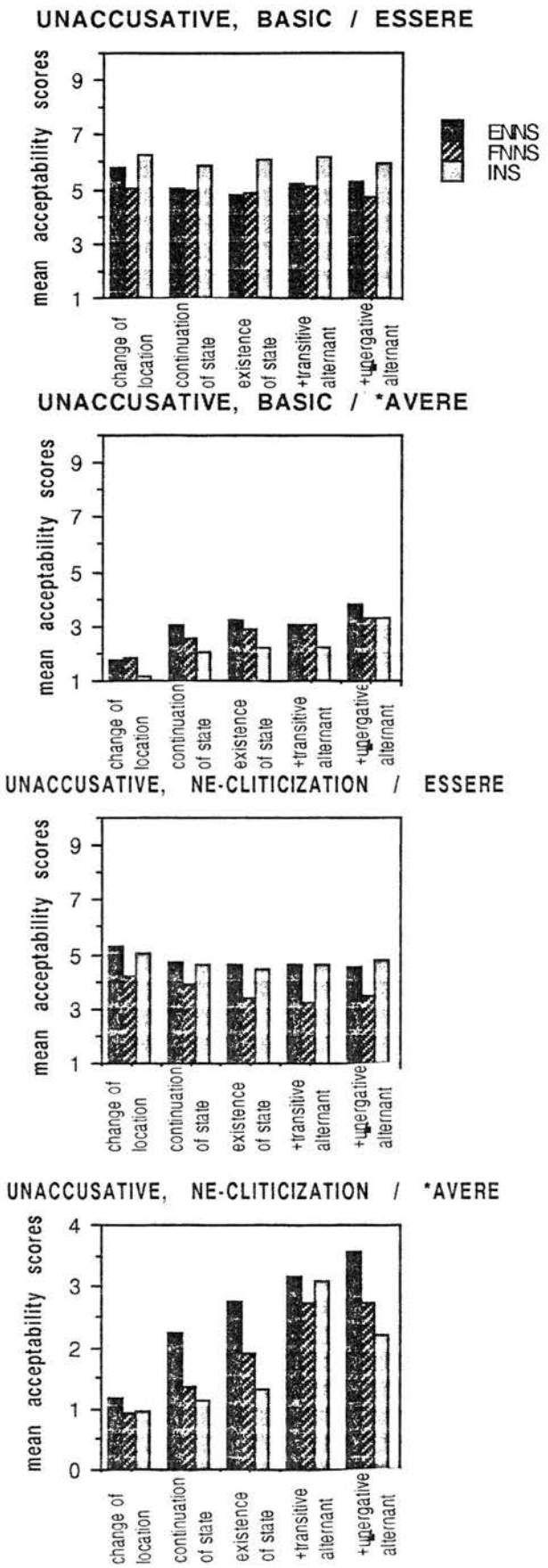
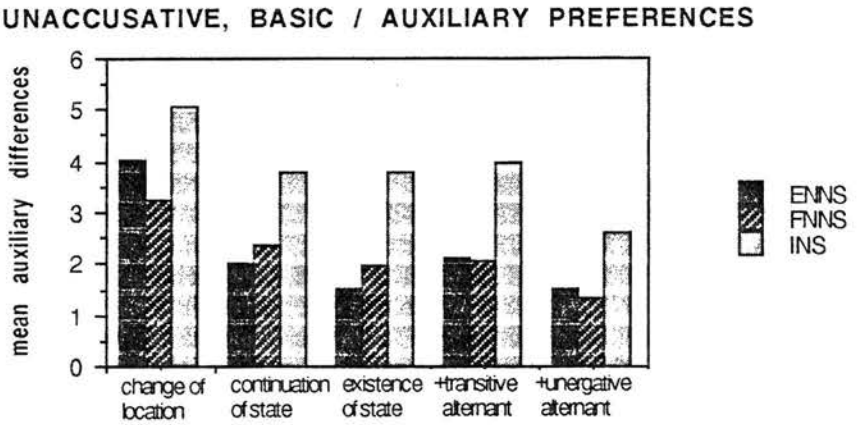


FIGURE C.11: Mean auxiliary preferences for **unaccusative** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).



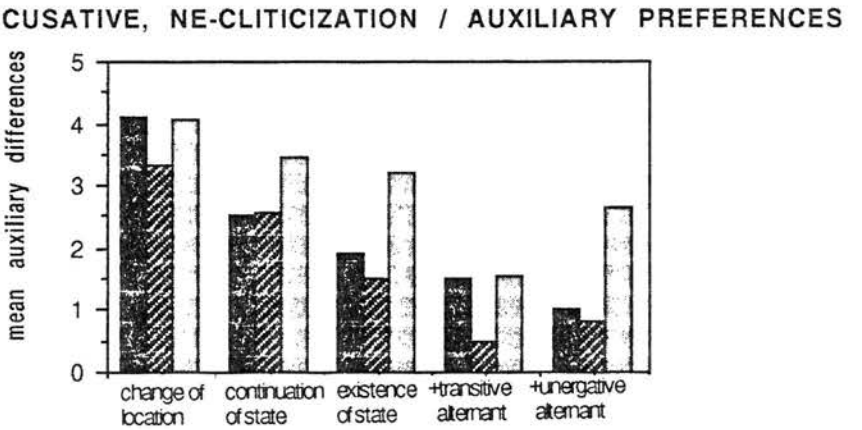


TABLE C.11: Mean acceptability judgments on **restructuring** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).

	ENNS	FNNS	INS
RAISING, BASIC, NON-RESTRUCTURED	5.042	4.650	5.167
RAISING, CLEFTING, NON-RESTRUCTURED	2.417	2.100	3.333
RAISING, CLIT-CLIMB, NON-RESTRUCTURED	3.792	3.300	3.750
RAISING, NO-CLIT-MOV, NON-RESTRUCTURED	3.625	4.100	4.389
RAISING, BASIC, RESTRUCTURED	4.042	4.000	4.944
RAISING, CLEFTING, RESTRUCTURED	2.083	1.800	2.444
RAISING, CLIT-CLIMB, RESTRUCTURED	4.125	4.050	4.667
RAISING, NO-CLIT-MOV, RESTRUCTURED	4.083	4.250	4.306
CONTROL, BASIC, NON-RESTRUCTURED	5.375	5.000	5.778
CONTROL, CLEFTING, NON-RESTRUCTURED	2.958	2.300	3.833
CONTROL, CLIT-CLIMB, NON-RESTRUCTURED	2.875	2.700	2.778
CONTROL, NO-CLIT-MOV, NON-RESTRUCTURED	4.750	4.350	4.889
CONTROL, BASIC, RESTRUCTURED	3.375	4.100	4.417
CONTROL, CLEFTING, RESTRUCTURED	2.167	1.900	2.194
CONTROL, CLIT-CLIMB, RESTRUCTURED	3.500	4.050	4.611
CONTROL, NO-CLIT-MOV, RESTRUCTURED	3.000	3.650	3.861

TABLE C.12: Mean auxiliary preferences for **restructuring** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).

	ENNS	FNNS	INS
RAISING, BASIC	1.000	0.650	0.222
RAISING, CLEFTING	0.333	0.300	0.889
RAISING, CLITIC-CLIMBING	0.125	-0.150	0.917
RAISING, NO-CLITIC-MOVEMENT	0.042	-0.700	0.083
CONTROL, BASIC	2.000	0.900	1.361
CONTROL, CLEFTING	0.792	0.400	1.639
CONTROL, CLITIC-CLIMBING	-0.542	0.100	1.167
CONTROL, NO-CLITIC-MOVEMENT	1.750	0.700	1.028

FIGURE C.12: Mean acceptability judgments on **restructuring (Raising)** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).

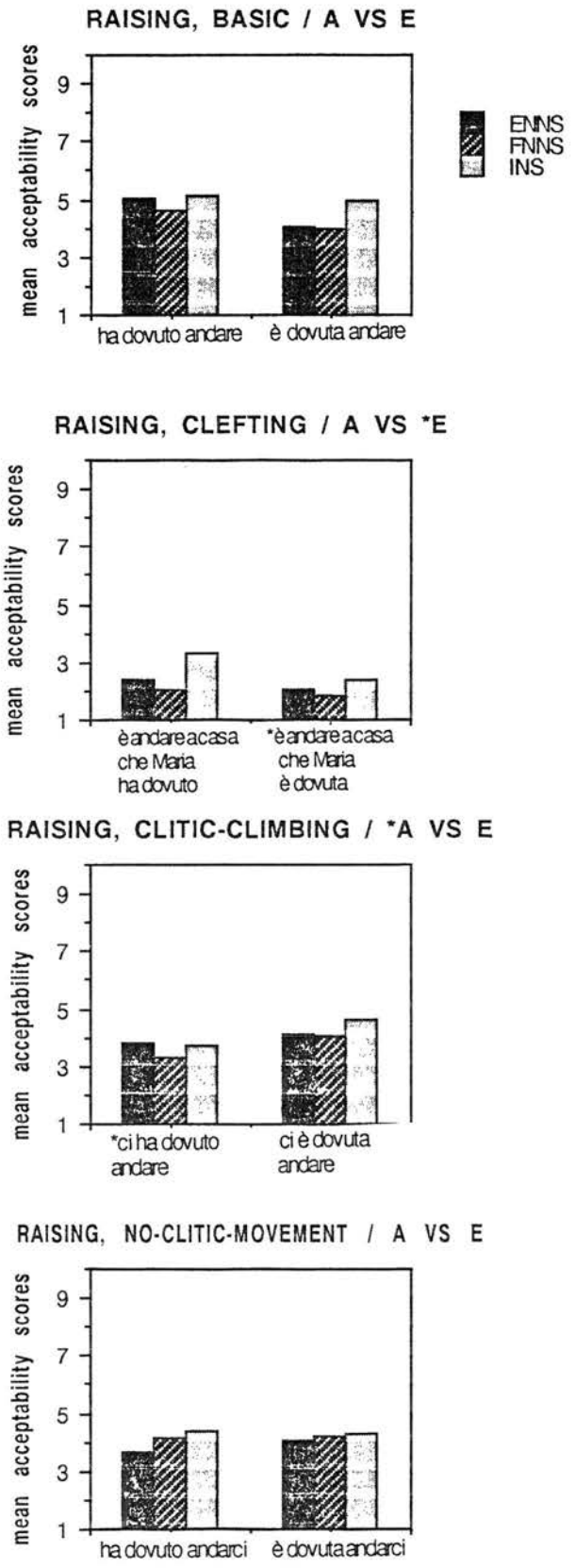
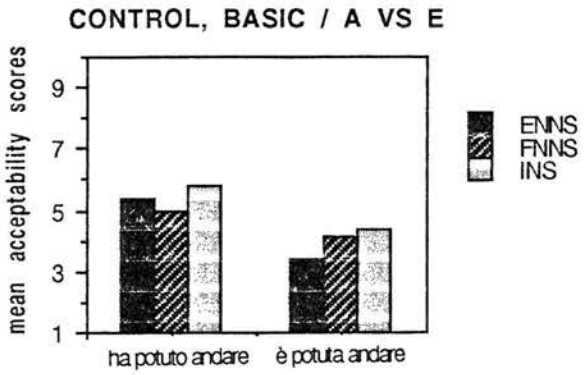
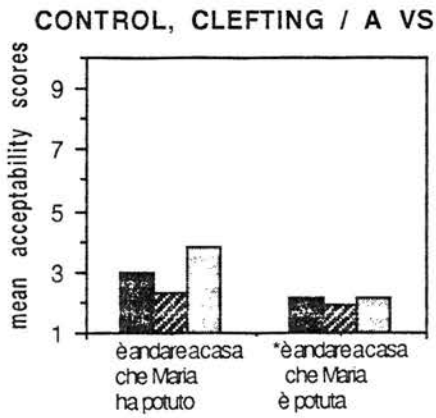
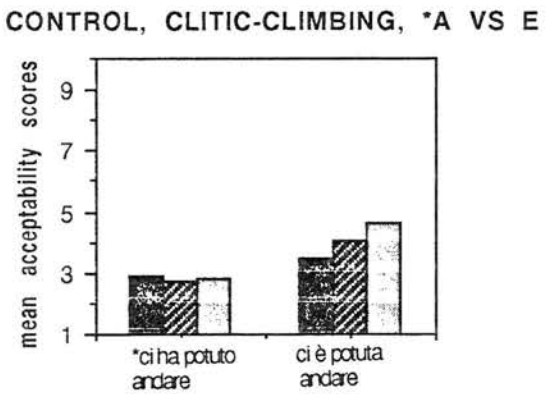


FIGURE C.13: Mean acceptability judgments on **restructuring (Control)** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).







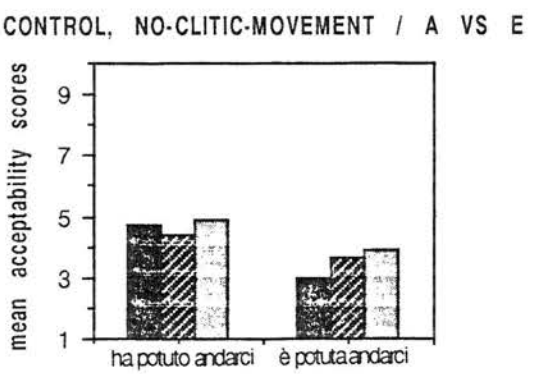


FIGURE C.14: Mean auxiliary preferences for **restructuring** verbs (ENNS = English near-native speakers, FNNS = French near-native speakers, INS = Italian native speakers).